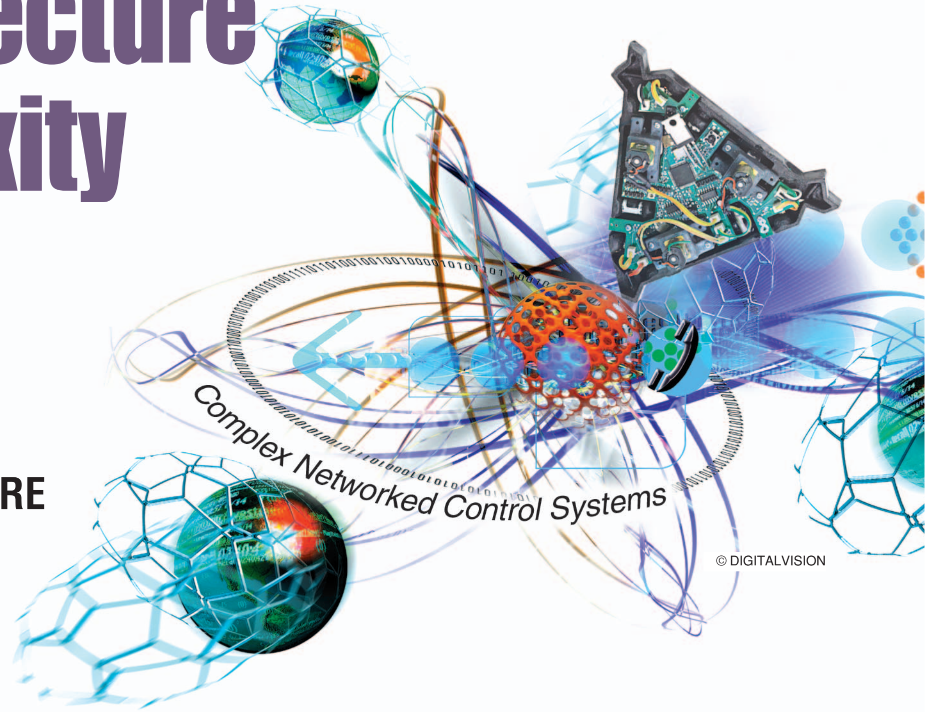
The Architecture of Complexity

ALBERT-LÁSZLÓ BARABÁSI

FROM NETWORK STRUCTURE TO HUMAN DYNAMICS



We are surrounded by complex systems, from cells made of thousands of molecules to society, a collection of billions of interacting individuals. These systems display signatures of order and self-organization. Understanding and quantifying this complexity is a grand challenge for science. Kinetic theory, developed at the end of the 19th century, shows that the measurable properties of gases, from pressure to temperature, can be reduced to the random motion of atoms and molecules. In the 1960s and 1970s, researchers developed systematic approaches to quantifying the transition from disorder to order in material systems such as magnets and liquids. Chaos theory dominated the quest to understand complex behavior in the 1980s with the message that unpredictable behavior can emerge from the nonlinear interactions of a few components. The 1990s was the decade of fractals, quantifying the geometry of patterns emerging in self-organized systems, from leaves to snowflakes.

Despite these conceptual advances, a complete theory of complexity does not yet exist. When trying to characterize complex systems, the available tools fail for various reasons. First, most complex systems are not made of identical components, such as gases and magnets. Rather, each gene in a cell or each individual in society has its own characteristic behavior. Second, while the interactions among the components are manifestly nonlinear, truly chaotic behavior is more the exception than the rule. Third, and most important, molecules and people do not obey either the extreme disorder of gases, where any molecule can collide with any other molecule, or the extreme order of magnets, where spins interact only with their immediate neighbors in a periodic lattice. Rather, in complex systems, the interactions form *networks,* where each node interacts with only a small number of selected partners whose presence and effects might be felt by far away nodes.

Networks exist everywhere and at every scale. The brain is a network of nerve cells connected by axons, while cells are networks of molecules connected by biochemical reactions. Societies, too, are networks of people linked by friendship, family, and professional ties. On a larger scale, food webs and ecosystems can be represented as networks of species. Furthermore, networks pervade technology; examples include the Internet, power grids, and transportation systems. Even the language used to convey thought is a network of words connected by syntactic relationships.

Despite the pervasiveness of networks, however, their structure and properties are not yet fully understood. For example, the mechanisms by which malfunctioning genes in a complex genetic network lead to cancer are not obvious, and rapid diffusion through social and communications networks that lead to epidemics of diseases and computer viruses, is not well characterized. Moreover, it is important to understand how some networks continue to function despite the failure of a majority of their nodes.

Recent research is beginning to answer such questions [1]–[6]. Over the past few years, scientists have discovered that complex networks have an underlying architecture guided by universal principles. For instance, many networks, from the World Wide Web (WWW) to the cell's metabolic system to the actors of Hollywood, are dominated

1066-033X/07/$25.00©2007IEEE

by a small number of nodes that are highly connected to other nodes. These important nodes, called *hubs*, greatly affect a network's overall behavior. As described in this article, hubs make the network robust against accidental failures but vulnerable to coordinated attacks.

The purpose of this article is to illustrate, through the example of human dynamics, that a thorough understanding of complex systems requires an understanding of network dynamics as well as network topology and architecture. After an overview of the topology of complex networks, such as the Internet and the WWW, data-driven models for human dynamics are given. These models motivate the study of network dynamics and suggest that complexity theory must incorporate the interactions between dynamics and structure. The article also advances the notion that an understanding of network dynamics is facilitated by the availability of large data sets and analysis tools gained from the study of network structure.

THE RANDOM NETWORK PARADIGM

Complex networks were originally thought of as being completely random. This paradigm has its roots in the work of Paul Erdős and Alfréd Rényi who, aiming to describe networks in communications and life sciences, suggested in 1959 that networks be modeled as random graphs [7], [8]. Their approach takes N nodes and connects them by L randomly placed links. The simplicity of the model and the elegance of the theory led to the emergence of random networks as a mathematical field of study [7]–[9].

A key prediction of random network theory is that, despite the random placement of links, most nodes are assigned approximately the same number of links. Indeed, in a random network the nodes follow a bell-shaped Poisson distribution. Finding nodes that have a significantly greater or smaller number of links than a randomly chosen node is therefore rare. Random networks are also called exponential networks because the

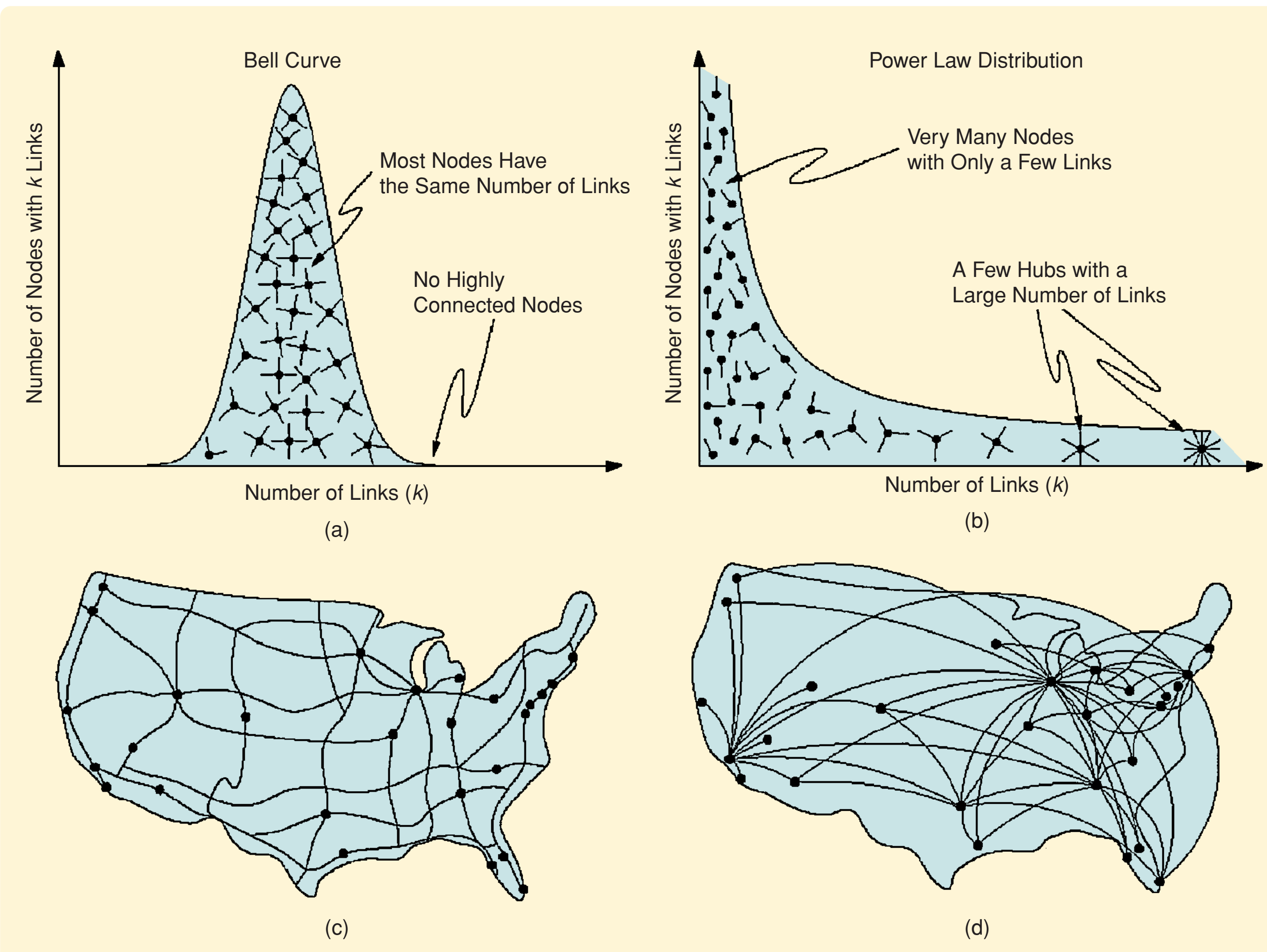


FIGURE 1 Random and scale-free networks. The degree distribution of a random network follows a Poisson distribution close in shape to the bell curve, telling us that most nodes have the same number of links, and that nodes with a large number of links don't exist (a). Thus, a random network is similar to a national highway network in which the nodes are the cities and the links are the major highways connecting them. Indeed, most cities are served by roughly the same number of highways (c). In contrast, the power-law degree distribution of a scale-free network predicts that most nodes have only a few links held together by a few highly connected hubs (b). Such a network is similar to the air traffic system, in which a large number of small airports are connected to each other by means of a few major hubs (d). After [1].

probability that a node is connected to k other nodes decreases exponentially for large k (Figure 1). The Erdős-Rényi model, however, raises the question as to whether networks observed in nature are truly random. Could the Internet, for example, offer fast and seamless service if computers were randomly connected to each other? Or could you read this article if the chemicals in your body suddenly decided to react randomly with each other, bypassing the rigid chemical web they normally obey? Intuitively the answer is no, since we suspect that behind each complex system there is an underlying network with nonrandom topology. The challenge of network structure studies, however, is to unearth the signatures of order from the collection of millions of nodes and links that form a complex network.

THE WORLD WIDE WEB AND INTERNET AS COMPLEX NETWORKS

The WWW contains more than a billion documents (Web pages), which represent the nodes of a complex network. These documents are connected by uniform resource locators (URLs), which are used to navigate from one document to another [Figure 2(a)]. To analyze the properties of the WWW, a map of how Web pages are linked to each other is obtained in [10] using a robot, or *Web crawler,* which starts from a given Web page and collects the page's outgoing links. The robot then follows each outgoing link to visit more pages, collecting their respective outgoing links [10]. Through this iterative process, a small but representative fraction of the WWW can be mapped out.

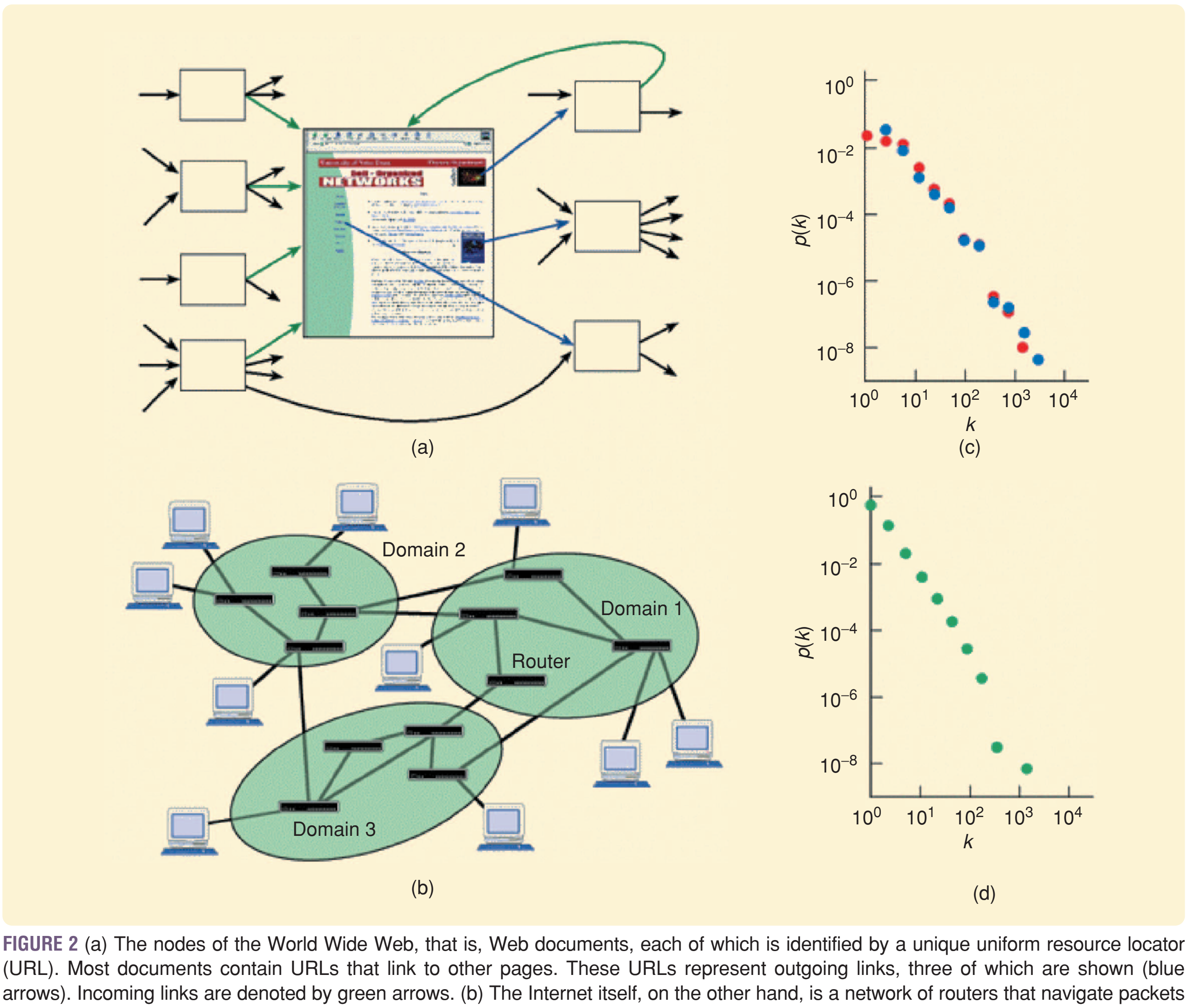


FIGURE 2 (a) The nodes of the World Wide Web, that is, Web documents, each of which is identified by a unique uniform resource locator (URL). Most documents contain URLs that link to other pages. These URLs represent outgoing links, three of which are shown (blue arrows). Incoming links are denoted by green arrows. (b) The Internet itself, on the other hand, is a network of routers that navigate packets of data from one computer to another. The routers, which are connected to each other by physical or wireless links, are grouped into several domains. (c) The probability that a Web page has k_{in} (blue) or k_{out} (red) links follows a power law. The results are based on a sample of more than 325,000 Web pages. (d) The degree distribution of the Internet at the router level, where k denotes the number of links a router has to other routers. These results, which are based on more than 260,000 routers, demonstrate that the Internet exhibits power-law behavior. After [49].

Since the WWW is a directed network, each document is characterized by the number k_{out} of its outgoing links and the number k_{in} of its incoming links. The outgoing (incoming) degree distribution thus represents the probability *P(k)* that a randomly selected Web page has exactly k_{out} (k_{in}) links. Although random graph theory predicts that *P(k)* follows a Poisson distribution, the collected data indicate that *P(k)* follows the power-law distribution shown in Figure 2(c) and described by

$$P(k) \sim k^{-\gamma}, \tag{1}$$

where $\gamma_{out} \cong 2.45$ ($\gamma_{in} \cong 2.1$).

As illustrated in Figure 1, major topological differences exist between a network with a Poisson connectivity distribution and one with a power-law connectivity distribution. Indeed, most nodes in an undirected random network have approximately the same number of links given by $k \approx \langle k \rangle$, where $< k >$ represents the average degree. The exponential decay of the Poisson distribution *P(k)* guarantees the absence of nodes with significantly more links than $\langle k \rangle$ and thus imposes a natural scale in the network. In contrast, the power-law distribution implies that nodes with few links are abundant, while a small number of nodes have a large number of links. A map of the U.S. highway system, where cities are nodes and highways are links, illustrates an exponential network. Most cities are located at the intersection of two to five highways. On the other hand, a scale-free network is similar to the airline routing maps displayed in flight magazines. While most airports are served by few carriers, a few hubs, such as Chicago or Frankfurt, have links to almost all other U.S. or European airports, respectively. Thus, just like the smaller airports, the majority of WWW documents have a small number of links, and, while these links are not sufficient by themselves to ensure that the network is fully connected, the few highly connected hubs guarantee that the WWW is held together.

Unlike Poisson distributions, a power-law distribution does not possess an intrinsic scale, and its average degree $< k >$ does not convey much information about the network structure. The absence of an intrinsic scale in k in networks with a power-law degree distribution motivates the concept of a scale-free network [11]. A scale-free network is therefore a network with a degree distribution that obeys a power law. Empirical measurements, however, indicate that real networks deviate from simple power-law behavior. The most typical deviation is the flattening of the degree distribution at small values of k, while a less typical deviation is the exponential cutoff for high values of k. Thus, a proper fit to the degree distribution of real networks has the form $P(k) \sim (k + k_0)^{-\gamma} \exp(-k/k,)$, where k_0 is the small-degree cutoff and k_x is the length scale of the high-degree exponential cutoff. The scale-free behavior of real networks is therefore evident only between k_0 and k_x.

The scale-free topology of the WWW motivates the search for inhomogeneous topologies in other complex systems such as the Internet. Unlike the WWW, the Internet is a physical network whose nodes are routers and domains and whose links are the phone lines and optical cables that connect the nodes [Figure 2(b)]. Due to its physical nature, the Internet is expected to be structurally different from the WWW, where adding a link to an arbitrary remote Web page is as easy as linking to a Web page on a computer in the next room. The Internet network, however, also appears to follow a power-law degree distribution as observed in [12] [see Figure 2(b)]. In particular, the degree distribution is shown to follow a power law with an exponent $\gamma = 2.5$ for the router network and $\gamma = 2.2$ for the domain map, which indicates that the wiring of the Internet is also dominated by several highly connected hubs [12].

19 Degrees of Separation

Stanley Milgram showed empirically in 1967 that any two persons are typically five to six handshakes away from each other [13]. That is, most humans on Earth appear to live in a *small world*. This feature of social networks is known as the *six-degrees of separation* property [14]. In addition, sociologists repeatedly argue that nodes in social networks are grouped in small clusters. These clusters represent circles of friends and acquaintances, and, within each cluster, a node is connected to all other nodes but has only sparse links to the outside world [15]. The question then arises as to whether the small world model is applicable to the WWW and the Internet.

Since a complete map of the WWW is not available [16], small computer models of the WWW are used in [10], where the link distribution matches the measured functional form and where the shortest distances between any two nodes are identified and averaged over all node pairs to obtain the average node separation d. By repeating this process for networks of different sizes using *finite size scaling*, a standard procedure of statistical mechanics, it is inferred in [10] that $d = 0.35 + 2.06. \log(N)$, where N is the number of WWW nodes. For the 800 million nodes of the WWW in 1999, the typical shortest path between two randomly selected pages is thus around 19, assuming that such a path exists, which is not always guaranteed because of the Web's directed nature. As shown empirically in [17], however, for 200 million nodes this distance is 16, in contrast to 17 as predicted in [10].

These results indicate that the WWW represents a small world and that the typical number of clicks between two Web pages is around 19, despite the current number of more than 1 billion online pages. Moreover, the WWW displays a high degree of clustering [18], that is, the probability that two neighbors of a given node are also linked is much greater than the value expected for a random network. Finally, results reported in [1] indicate that the Internet also possesses a small-world structure.

EVOLVING NETWORKS

The emergence of scale-free structures and the power-law degree distribution can be traced back to two mechanisms. These mechanisms are absent from the classical random graph models although they are present in various complex networks [11]. First, traditional graph-theoretic models assume that the number of nodes in a network is fixed. In contrast, the WWW continuously expands by adding new Web pages, while the Internet grows with the installation of new routers, computers, and communication links. Second, while random graph models assume that the links are randomly distributed, most real networks exhibit *preferential attachment*. Indeed, a person is more likely to link a Web page to highly connected documents on the WWW, whose existence is well known. Network engineers also tend to connect their institution's computers to the Internet through high-bandwidth nodes, which inevitably attract a large number of other consumers and links.

Based on the increasing number of nodes as well as on preferential attachment, a simple model in which a new node is added to the network at each time step is considered in [11]. The new node is then linked to some of the nodes already present in the system (Figure 3). The probability $\Pi(k)$ that a new node connects to a node with k links follows a preferential attachment rule such as

$$\Pi(k) = \frac{k}{\Sigma_i k_i}, \tag{2}$$

where the sum is over all nodes in the network. Numerical simulations indicate that the resulting network is scale free, and the probability that a node has k links follows (1) with exponent $\gamma = 3$ [11]. The power-law nature of the distribution is predicted by a rate-equation-based approach [19] as well as from an exact solution of the scale-free model [20]. This simple model illustrates how growth and preferential attachment jointly lead to the appearance of the hub hierarchy that exemplifies the scale-free structure. A node with more links increases its connectivity faster than nodes with fewer links, since incoming nodes tend to connect to it with higher probability as described in (2). This model leads to a *rich-get-richer* positive-feedback phenomenon, which is evident in some competitive systems.

THE ACHILLES' HEEL OF THE INTERNET

As the world economy becomes increasingly dependent on the Internet, a concern arises about whether the Internet's functionality can be maintained under failures and hacker attacks. The Internet has so far proven remarkably resilient against failures. Even though around 3% of the routers are typically down at a particular moment, we rarely observe major Internet disruptions. How did the Internet come to be so robust? While significant error tolerance is built into the protocols that govern packet-switching communications, the scale-free topology of the Internet also plays a crucial role in making it more robust.

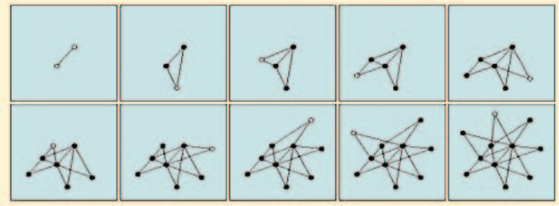


FIGURE 3 Birth of a scale-free network. The scale-free topology is a natural consequence of the ever-expanding nature of real networks. Starting from two connected nodes (top left), in each panel a new node, which is shown as an open dot, is added to the network. When deciding where to link, new nodes prefer to attach to the more connected nodes. Thanks to growth and preferential attachment, a few highly connected hubs emerge. After [1].

Percolation concepts provide one approach to understanding the scale-free induced error tolerance of the Internet. Percolation theory specifies that the random removal of nodes from a network results in an inverse percolation transition. When a critical fraction f_c of nodes is removed, the network fragments into tiny, noncommunicating islands of nodes. However, simulations of scale-free networks do not support percolation's theory prediction [21]. With up to 80% of the nodes of a scale-free network removed, the remaining nodes remain part of a compact cluster. The disagreement is resolved in [22] and [23], where it is shown that as long as the connectivity exponent γ in (1) is smaller than three, which is the case for most real networks, including the Internet, the critical threshold for fragmentation is $f_c = 1$. This result demonstrates that scale-free networks cannot be broken into pieces by the random removal of nodes. This extreme robustness relative to random failures is rooted in the inhomogeneous network topology. Since there are far more weakly connected nodes than hubs, random removal most likely affects the less-connected nodes. The removal of a node with a small degree does not significantly disrupt the network topology, just as the closure of a local airport has little impact on international air traffic.

Their inhomogeneous topology, however, makes scale-free networks especially vulnerable to targeted attacks [21]. Indeed, the removal of a small fraction of the most connected nodes (hubs) might break the network into pieces. These findings illustrate the underlying topological vulnerability of scale-free networks. In fact, while the Internet is not expected to break under the random failure of routers and links, well-informed hackers can easily handicap the network by targeting hubs for attacks.

While error tolerance and vulnerability to attacks are consequences of the scale-free property, the reverse is not necessarily true. Networks that are resilient relative to random attacks but that fragment under targeted attacks are not necessarily scale free. For example, the hub-and-spoke network, in which all nodes connect to a central node, is the most resilient network relative to random failures.

> **Stanley Milgram showed empirically in 1967 that any two persons are typically five to six handshakes away from each other.**

Such a network fails only when the central hub is removed, an event whose probability of occurrence is $1/N$ for a network with N nodes. Therefore, it is more appropriate to define a scale-free network based on its degree distribution rather than its robustness properties.

SCALE-FREE EPIDEMICS

The structure of scale-free networks can help explain the spread of computer viruses, diseases, and fads. Diffusion theories by both epidemiologists and marketing experts predict the presence of a *critical threshold* for successful propagation throughout a population or a network. A virus that is less virulent or a fad that is less contagious than the critical threshold inevitably dies out, while those above the threshold multiply exponentially, eventually penetrating the entire network.

As shown in [24], however, the critical threshold of a scale-free network is zero. Therefore, all viruses, even those that are only weakly contagious, eventually spread and persist in the system. This result explains why *Love Bug*, the most damaging computer virus thus far, remains the seventh most prevalent virus even a year after its supposed eradication. Hubs are again the key to this surprising behavior. In fact, since hubs are highly connected, at least one of them is likely to become infected by a single corrupted node. Moreover, once a hub is infected, it broadcasts the virus to numerous other nodes, eventually compromising other hubs that help spread the virus throughout the system.

Because biological viruses spread on scale-free social networks, scientists need to consider the interplay between network topology and epidemics. Specifically, in a scale-free contact network, the traditional public-health approach of random immunization can fail by missing some of the hubs. The topology of scale-free networks suggests an alternative approach. By targeting the hubs, that is, the most connected individuals, the immunizations become effective after reaching only a small fraction of the population [25]–[27]. Identifying the hubs in a social network, however, is much more difficult than in other types of systems such as the Internet. But when a small fraction of the random acquaintances of randomly selected individuals is immunized, the hubs are highly likely to be immunized since hubs are acquainted with many people [26].

HUMAN DYNAMICS AND THE TEMPORAL BEHAVIOR OF SINGLE NODES

The above discussion focuses on one aspect of complex networks, namely, their topology. We now consider the role of *network dynamics*. Indeed, most complex systems of practical interest, from the cell to the Internet to social networks, are fascinating because of their temporal behavior. While such systems have nontrivial network topologies, the role of their topology is to serve as a skeleton on which dynamical processes, from information to material transfer, take place. Topological network theory, while indispensable in describing these dynamical processes, does not yet fully account for the complex behavior displayed by these systems. There is thus a need to characterize the dynamical processes taking place on complex networks and to understand the interplay between topology and dynamics.

Next, we describe recent advances in the quantitative understanding of human dynamics, since the dynamics of many social, technological, and economic networks are driven by individual human actions.

Current models of human dynamics in areas such as risk assessment and communications assume that human actions are randomly distributed in time and are well approximated by Poisson processes [28]–[30]. In the following, evidence is presented to show that the timing of many human activities, ranging from communication to entertainment and work patterns, follow non-Poisson statistics [31]–[42]. These statistics are characterized by bursts of rapidly occurring events separated by long periods of inactivity. This *bursty* nature of human behavior is a consequence of a decision-based queuing process. When individuals execute tasks based on a perceived priority, the tasks' timing follows a heavy-tailed distribution, with most tasks being rapidly executed while a few tasks experience long waiting times [31]. In contrast, priority-blind execution is well approximated by uniform inter-event statistics.

Humans participate in a large number of distinct daily activities. These activities range from electronic communication to browsing the Web to initiating financial transactions and engaging in entertainment and sports. Factors ranging from work and sleep patterns to resource availability and interaction with other individuals determine the timing of each daily human activity. Finding regularities in human dynamics, apart from the obvious daily and seasonal periodicities, appears difficult if not impossible. A quantitative understanding of network dynamics driven by human activities might therefore appear hopeless. We show next that this appearance is not entirely accurate and that human dynamics are driven by interesting, reproducible mechanisms.

Poisson processes assume that, in a time interval of duration dt, an individual engages in a specific action with probability qdt, where q is the frequency of the monitored activity. According to such models, the time interval between two consecutive actions by the same individual, called the waiting or inter-event time, follows an exponential distribution (Figure 4) [28], [42]. Poisson processes are widely used to quantify the consequences of human actions, such as modeling traffic flow patterns or frequency of accidents [28], call center staffing [29], and inventory control [30], or to estimate the number of congestion-caused blocked calls in mobile communications [32]. Measurements indicate, however, that the timing of many human actions systematically deviates from the Poisson-based prediction. In fact, waiting or inter-event times are better approximated by a heavy-tailed, Pareto distribution (Figure 4). The differences between Poisson and heavy-tailed behavior are striking. A Poisson distribution decreases exponentially, forcing the consecutive events to follow each other at regular time intervals and forbidding

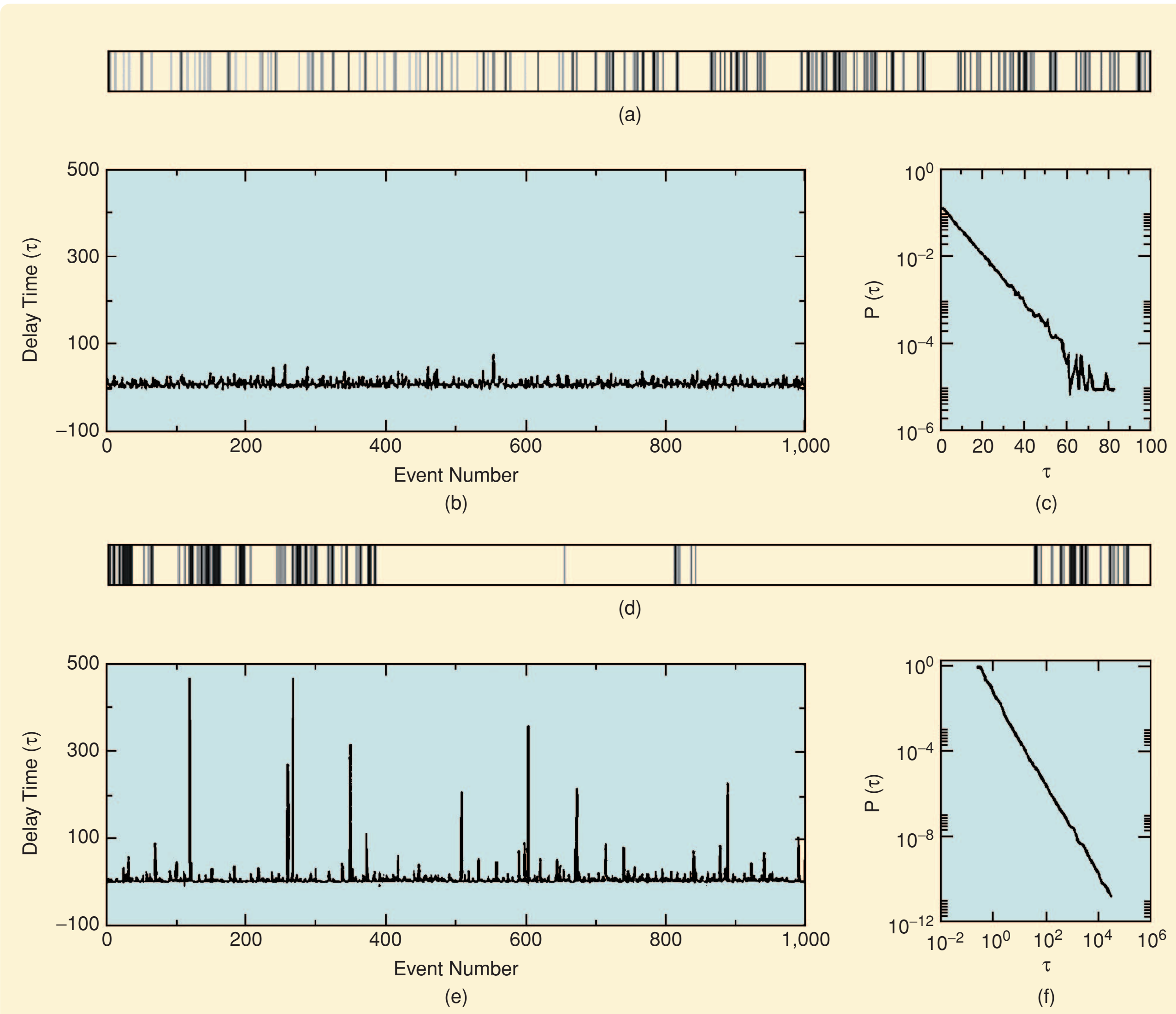


FIGURE 4 The difference between the activity patterns predicted by a Poisson process (top) and the heavy tailed distributions observed in human dynamics (bottom). (a) Succession of events predicted by a Poisson process, which assumes that in any moment an event takes place with probability q. The horizontal axis denotes time, each vertical line corresponding to an individual event. Note that the inter-event times are comparable to each other, long delays being virtually absent. (b) The absence of long delays is visible on the plot showing the delay times τ for 1,000 consecutive events, the size of each vertical line corresponding to the gaps seen in (a). (c) The probability of finding exactly n events within a fixed time interval is $P(n;\ q) = e^{-qt}(qt)^n/n!$, which predicts that, for a Poisson process, the inter-event time distribution follows $P(\tau) = qe^{-q\tau}$, which is shown on a log-linear plot. (d) The succession of events for a heavy tailed distribution. (e) The waiting time τ of 1,000 consecutive events, where the mean event time is chosen to coincide with the mean event time of the Poisson process shown in (a)–(c). Note the large spikes in the plot, which correspond to long delay times. (b) and (e) have the same vertical scale, allowing comparison of the regularity of a Poisson process with the bursty nature of the heavy tailed process. (f) The delay time distribution $P(\tau) \sim \tau^{-2}$ for the heavy tailed process shown in (d) and (e) appears as a straight line with slope -2 on a log-log plot. After [42].

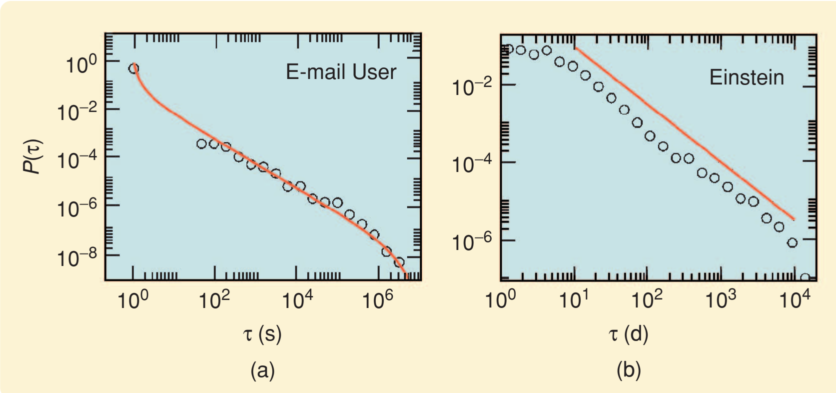


FIGURE 5 (a) The response time distribution of an e-mail user, where the response time is defined as the time interval between the time the user first sees an e-mail and the time the user sends a reply to it. The first symbol in the upper left corner corresponds to messages that are replied to right after the user notices them. The continuous line corresponds to the waiting time distribution of the tasks, as predicted by Model A discussed in the text, obtained for p= 0.999999 + 0.000005. (b) Distribution of the response times for the letters replied to by Einstein. The distribution is well approximated with a power-law tail with exponent $\alpha = 3/2$, as predicted by Model B. Note that while in most cases the identified reply is indeed a response to a received letter, there are exceptions as well since some of the much-delayed replies represent the renewal of a long-lost relationship. After [31].

long waiting times. In contrast, slowly decaying heavy-tailed processes allow for long periods of inactivity separating bursts of intense activity (Figure 4).

Two human activity patterns provide evidence for non-Poisson activity patterns in individual human behavior, namely, e-mail communication based on a data set capturing the sender, recipient, time, and size of each e-mail [33], [34], and the letter-based communication pattern of Einstein [43]. As Figure 5 shows, the response time distribution of both e-mails and letters is best approximated with

$$P(\tau) \sim \tau^{-\alpha}, \tag{3}$$

where $\alpha = 1$ for e-mail communications and $\alpha = 3/2$ for letters. These results indicate that an individual's communication pattern has a bursty non-Poisson character. For a short time period, a user sends out several e-mails or letters in quick succession followed by long periods of no communication.

Measurements capturing the distribution of the time differences between consecutive instant messages sent by individuals during online chats [35] show a similar pattern. Professional tasks, such as the timing of job submissions on a supercomputer [36], directory listings, and file transfer (FTP) requests initiated by individual users [37], the timing of print jobs submitted by users [38], or the return visits of individual users to a Web site [39], are also reported to display non-Poisson features. Similar patterns emerge in economic transactions that describe the number of hourly trades in a given security [39] or the time interval distribution between individual trades in currency futures [40]. Finally, heavy-tailed distributions characterize entertainment events such as the time intervals between consecutive online games played by the same user [41].

The wide range of human activity patterns that follow non-Poisson statistics suggests that the observed bursty character of such patterns reflects a fundamental and potentially generic feature of human dynamics. Next, we show that this phenomenon is a consequence of a queuing process driven by human decision making. Whenever an individual is presented with multiple tasks and is asked to choose among them based on a perceived priority parameter, the waiting time of the various tasks is Pareto distributed [31], [42]. In contrast, first-come, first-served as well as random task executions, both of which are typical in service-oriented or computer-driven environments, lead to Poisson-like dynamics.

Most human-initiated events require an individual to prioritize various activities. At the end of each activity, an individual decides what to do next, such as send an e-mail, shop, or place a phone call, thus allocating time and resources for the chosen activity. Consider an individual operating with a priority list of L tasks. A task is removed from the list once executed, and new tasks are added to the list as soon as they emerge. To compare the urgency of the various tasks on the list, the agent or individual assigns a priority parameter x to each task. The waiting time of a given task before it is executed depends on the method used by the agent to choose the order of task execution. In this respect, the following three selection protocols are relevant for human dynamics:

i) The simplest protocol is the first-in, first-out protocol, which executes tasks in the order that the tasks are added to the list. This protocol is typical in service-oriented processes, such as the execution of orders in a restaurant or in directory assistance and consumer-support applications. A task remains on the list before it is executed for as long as it takes to perform all tasks that are ahead of it. When the time required to perform a task is chosen from a bounded distribution (the second moment of the distribution is finite), the waiting-time distribution develops an exponential tail, indicating that most tasks experience approximately the same waiting time.
ii) In the second protocol, tasks are executed in a random order, irrespective of their priority or time spent on the list. This protocol is typical, for example, in educational settings where students are called on randomly

as well as in some Internet packet-routing protocols. In this case, the waiting-time distribution of individual tasks is also exponential.

iii) In most human-initiated activities, task selection is not random. Individuals usually execute the highest priority items on their list. The resulting execution dynamic is different from i) and ii). High-priority tasks are executed soon after they are added to the list, while low-priority items must wait until all higher priority tasks are cleared. This protocol forces lower priority items to stay on the list longer than higher priority ones.

In the following, we show that priority selection, practiced by humans on a daily basis, is a likely source of the heavy tails observed in human-initiated processes. We consider two models for this protocol.

Model A [42]

Assume that an individual has a priority list with L tasks and that each task is assigned a priority parameter x_i, $i = 1, \ldots, L$, chosen from a distribution $\rho(x)$. At each time step, the individual selects the highest priority task from the list and executes it, thus removing it from the list. At that moment, a new task is added to the list, its priority again chosen from $\rho(x)$. This simple model ignores the possibility that the individual occasionally selects a low-priority item for execution before all higher priority items are completed, a situation often seen for tasks with deadlines. The deadline-driven execution may be incorporated in this scenario by assuming that the agent executes the highest priority item with probability p, then executes with probability $1 - p$ a randomly selected task, independent of its priority. The limit $p \to 1$ therefore describes the deterministic iii) protocol, when the highest priority task is always chosen for execution, while corresponds to the random-choice protocol discussed in ii).

Model B [43], [44]

Assume that tasks arrive to the priority list at the rate λ following a Poisson process with exponential arrival-time distribution. The arrival of each new task increases the length of the list from L to $L + 1$. The tasks are executed at the rate μ, reflecting the overall time an individual devotes to its priority list. Once a task is executed, the length of the priority list decreases from L to $L - 1$. Each task is assigned a discrete priority parameter x_i upon arrival, such that the highest-priority unanswered task is always chosen for execution. The lowest priority task must wait the longest before execution, therefore dominating the waiting-time probability density for large waiting times.

The only difference between model A and model B is in the length of the queue L. In model A, the queue length is fixed and remains unchanged during the model's evolution, while, in model B, the queue length L can fluctuate as tasks arrive or are executed. This small difference, however, has a large impact on the distribution of the waiting times of tasks on the priority list. Indeed, numerical and analytical results indicate that both models give rise to a power-law waiting-time distribution. Model A predicts that, in the limit $p \to 0$, the waiting times follow (4) with $\alpha = 1$ [6], [42], which agrees with the observed scaling for e-mail communications, Web browsing, and several other human activity patterns. In contrast, for model B, the waiting-time distribution follows (3), with exponent $\alpha = 3/2$ as long as $\mu \leq \lambda$, which agrees with the results obtained for the mail correspondence patterns of Einstein [31], [43], [44].

These results indicate that human dynamics are described by at least two universality classes characterized by empirically distinguishable exponents. In searching to explain the observed heavy-tailed human activity patterns, only single queues are considered. In reality, actions are not performed independently since most daily activities are embedded in a web of actions involving other individuals. Indeed, the timing of an e-mail sent to user A may well depend on the time user A receives an e-mail from user B. A future goal of human dynamics research is to understand how various human activities and their timing are affected by the fact that individuals are embedded in a network environment. Such understanding can bring together the study of network topology and network dynamics.

FROM NETWORK THEORY TO A THEORY OF COMPLEXITY

As it stands today, network theory is not a proxy for a theory of complexity. Network theory currently addresses the emergence and structural evolution of the skeleton of a complex system. The ultimate understanding of the overall behavior of a complex network must, however, account for its architecture as well as the nature of dynamical processes taking place on such a network. Therefore, structural network theory is by no means the end of a journey but rather an unavoidable step toward the ultimate goal of understanding complex systems. Should a theory of complexity ever be completed, it must incorporate the newly discovered fundamental laws governing the architecture of complex systems [45].

At the nodes and links level, each network is driven by apparently random and unpredictable events. Despite this microscopic randomness, however, a few fundamental laws and organizing principles are helping to explain the topological features of such diverse systems as the cell, the Internet, and society. This new type of universality is driving the multidisciplinary explosion in network science. As shown in this article, results indicating some degree of universality in the behavior of individual nodes in complex systems are emerging. So far, these results refer to human dynamics only, although similar results have emerged lately on the nature of fluctuations in complex networks [46]–[48].

ACKNOWLEDGMENTS

This research was supported by NSF ITR 0426737, NSF ACT/SGER 0441089, the NSF NetWorkBench awards, and an award from the James S. McDonnell Foundation. I wish to also thank Alexei Vazquez and Suzanne Aleva for their assistance with the preparation of this article.

REFERENCES

[1] A.-L. Barabási, *Linked: The New Science of Networks*. Cambridge, MA: Plume Books, 2003.

[2] G. Caldarelli, *Scale-Free Networks*. London: Oxford Univ. Press, 2007.

[3] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, pp. 47–97, Jan. 2002.

[4] M.E.J. Newman, A.-L. Barabási, and D.J. Watts, *The Structure and Dynamics of Complex Networks*. Princeton, NJ: Princeton Univ. Press, 2006.

[5] S.N. Dorogovtsev and J.F.F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW.* New York: Oxford Univ. Press, 2003.

[6] S.H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, pp. 268–276, Mar. 2001.

[7] P. Erdös and A. Rényi, "On random graphs I," *Publ. Math. Debrecen*, vol. 6, pp. 290–297, 1959.

[8] P. Erdös and A. Rényi, "On the evolution of random graph," *Publ. Math. Inst. Hungarian Acad. Sci.* vol. 5, pp. 17–61, 1960.

[9] B. Bollobás, *Random Graphs*. New York: Academic, 1985.

[10] R. Albert, H. Jeong, and A.-L. Barabási, "Diameter of the World Wide Web," *Nature*, vol. 401, pp. 130–131, Sept. 1999.

[11] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science,* vol. 286, pp. 509–512, Oct. 1999.

[12] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," in *Proc. ACM SIGCOMM 99*, 1999, pp. 251–262.

[13] S. Milgram, "The small world problem," *Psychol. Today*, vol. 1, pp. 60–67, May 1967.

[14] J. Guare, *Six Degrees of Separation*. New York: Vintage Books, 1990.

[15] M.S. Granovetter, "The strength of weak ties," *Amer. J. Sociol.*, vol. 78, no. 6, pp. 1360–1380, 1973.

[16] S. Lawrence and C.L. Giles, "Searching the World Wide Web," *Science*, vol. 280, pp. 98–100, Apr. 1998.

[17] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajalopagan, R. Stata, A. Tomkins, and J. Weiner, "Graph structure in the Web," in *Proc. 9th Int. World Wide Web Conf. Computer Networks: Int. Journal Computer Telecommunications Networking*, vol. 33, 2000, pp. 309–320.

[18] B.A. Huberman and L. Adamic, "Growth dynamics of the World Wide Web," *Nature*, vol. 401, pp. 131, Sept. 1999.

[19] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, "Structure of growing networks with preferential linking," *Phys. Rev. Lett.*, vol. 85, pp. 4633–4636, Nov. 2000.

[20] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnady, "The degree sequence of a scale-free random graph process," *Random Structures Algorithms*, vol. 18, no. 3, pp. 279–290, Apr. 2001.

[21] R. Albert, H. Jeong, and A.-L. Barabási, "The Internet's Achilles' heel: Error and attack tolerance in complex networks," *Nature*, vol. 406, pp. 378–382, July 2000.

[22] R. Cohen, K. Reez, D. Ben-Avraham, and S. Havlin, "Resilience of the Internet to random breakdowns," *Phys. Rev. Lett.*, vol. 85, no. 21, pp. 4626–4628, 2000.

[23] R. Cohen, K. Reez, D. Ben-Avraham, and S. Havlin, "Breakdown of the Internet under intentional attack," *Phys. Rev. Lett.* vol. 86, no. 16, pp. 3682–3685, 2001.

[24] R. Pastor-Satorras and A. Vespignani, "Dynamical and correlation properties of the Internet," *Phys. Rev. Lett.,* vol. 87, no. 25, pp. 258701-1–4, 2001.

[25] Z. Dezsö and A.-L. Barabási, "Halting viruses in scale-free networks," *Phys. Rev. E,* vol. 65, pp. 055103-1–4, May 2002.

[26] S. Havlin and D. Ben-Avraham, "Efficient immunization strategies for computer networks and populations," *Phys. Rev. Lett.*, vol. 91, no. 24, 247901-1-4, 2003.

[27] R. Pastor-Satorras and A. Vespignani, "Immunization of complex networks," *Phys. Rev. E*, vol. 65, pp. 036104-1-8, Feb. 2002.

[28] F.A. Haight, *Handbook on the Poisson Distribution*. New York: Wiley, 1967.

[29] P. Reynolds, *Call Center Staffing*. Lebanon, TN: Call Center School Press, 2003.

[30] J.H. Greene, *Production and Inventory Control Handbook,* 3rd ed. NewYork: McGraw-Hill, 1997.

[31] A. Vázquez, J.G. Oliveira, Z. Dezsö, K.-I. Goh, I. Kondor, and A.-L. Barabási, "Modeling bursts and heavy-tails in human dynamics," *Phys. Rev. E,* vol. 73, no. 3, pp. 036127-1–19, 2006.

[32] H.R. Anderson, *Fixed Broadband Wireless System Design*. New York: Wiley, 2003.

[33] J.P. Eckmann, E. Moses, and D. Sergi, "Entropy of dialogues creates coherent structure in E-mail traffic," *Proc. Natl. Acad. Sci,* pp. 14333–14337, 2004.

[34] H. Ebel, L.I. Mielsch, and S. Bornholdt, "Scale-free topology of E-mail network," *Phys. Rev. E*, vol. 66, pp. 35103-1–4, Sept. 2002.

[35] C. Dewes, A. Wichmann, and A. Feldman, "An analysis of Internet chat systems," in *Proc. 2003 ACM SIGCOMM Conf. Internet Measurement (IMC-03)*, pp. 51–64.

[36] S.D. Kleban and S.H. Clearwater, "Hierarchical dynamics, interarrival times and performance," in *Proc. ACM/IEEE Supercomputing*, Phoenix, AZ, 2003, pp. 28–28.

[37] V. Paxson and S. Floyd, "Wide-area traffic: The failure of Poisson modeling," *IEEE/ACM Trans. Networking*, vol. 3, no. 3, pp. 226–244, 1995.

[38] U. Harder and M. Paczuski, "Correlated dynamics in human printing behavior," *Physica A*, vol. 361, no. 1, pp. 329–336, 2006.

[39] V. Plerou, P. Gopikirshnan, L.A.N. Amaral, X. Gabaix, and H.E. Stanley, "Economic fluctuations and anomalous diffusion," *Phys. Rev. E,* vol. 62, pp. 3023–3026, Sept. 2000.

[40] J. Masoliver, M. Montero, and G.H. Weiss, "Continuous-time random-walk model for financial distributions," *Phys. Rev. E*, vol. 67, pp. 021112/1–9, Feb. 2003.

[41] T. Henderson and S. Nhatti, "Modeling user behavior in networked games," in *Proc. ACM Multimedia*, Ottawa, Canada, 2001, pp. 212–220.

[42] A.-L. Barabási, "The origin of bursts and heavy tails in human dynamics," *Nature*, vol. 435, pp. 207–211, May 2005.

[43] J.G. Oliveira and A.-L. Barabási, "Human dynamics: The correspondence patterns of Darwin and Einstein," *Nature,* vol. 437, pp. 1251–1254, Oct. 2005.

[44] A. Cobham, "Priority assignment in waiting line problems," *J. Oper. Res. Soc. Amer*., vol. 2, pp. 70–76, Feb. 1954.

[45] A.-L. Barabási, "Taming complexity," *Nature Physics*, vol. 1, pp. 68–70, Nov. 2005.

[46] M.A. de Menezes and A.-L. Barabasi, "Fluctuations in network dynamics," Phys. Rev. Lett., vol. 92, no. 2, pp. 028701/1–4, 2004.

[47] M.A. de Menezes and A.-L. Barabási, "Separating internal and external dynamics of complex systems," *Phys. Rev. Lett*. vol. 93, no. 6, pp. 068701/1–4, 2004.

[48] J. Duch and A. Arenas, "Scaling of fluctuations in traffic on complex networks," *Phys. Rev. Lett.,* vol. 96, no. 21, art. no. 218702, pp. 218202/1-6, 2006.

[49] A.-L. Barabási, "The physics of the Web," *Physics World,* vol. 14, pp. 33–38, July 2001.

AUTHOR INFORMATION

Albert-László Barabási (alb@nd.edu) is the director of the Center for Complex Network Research, in the Department of Physics and Computer Science at the University of Notre Dame, Indiana. He studied physics and engineering at the University of Bucharest, obtained an M.Sc. in physics in 1991 from Eotvos Lorand University, Budapest, and a Ph.D. in physics in 1994 from Boston University. He is a fellow of the American Physical Society and a member of the Hungarian Academy of Sciences. He received the 2005 FEBS Anniversary Prize for Systems Biology and the 2006 Neumann Medal for Computer Science. He is the author of *Linked: The New Science of Networks*. He can be contacted at the University of Notre Dame, Department of Physics, 225 Nieuwland Science Hall, Notre Dame, IN 46556 USA.