PERSPECTIVE

Scale-Free Networks: A Decade and Beyond

Albert-László Barabási

For decades, we tacitly assumed that the components of such complex systems as the cell, the society, or the Internet are randomly wired together. In the past decade, an avalanche of research has shown that many real networks, independent of their age, function, and scope, converge to similar architectures, a universality that allowed researchers from different disciplines to embrace network theory as a common paradigm. The decade-old discovery of scale-free networks was one of those events that had helped catalyze the emergence of network science, a new research field with its distinct set of challenges and accomplishments.

Nature, society, and many technologies are sustained by numerous networks that are not only too important to fail but paradoxically for decades have also proved too complicated to understand. Simple models, like the one introduced in 1959 by mathematicians Pál Erdős and Alfréd Rényi (*1*), drove much of our thinking about interconnected systems. They assumed that complex systems are wired randomly together, a hypothesis that was adopted by sociology, biology, and computer science. It had considerable predictive power, explaining for example why everybody is only six handshakes from anybody else (*2*–*5*), a phenomenon observed as early as 1929 (*2*) but which resonated in physical sciences only after Duncan Watts and Stephen Strogatz extended its reach beyond sociology (*5*). Yet, the undeniable success of the random hypothesis did pose a fundamental question: Are real networks truly random? That is, could systems such as the cell or a society function seamlessly if their nodes, molecules, or people were wired randomly together? This question motivated our work as well, leading 10 years ago to the discovery of the scale-free property (*6*, *7*).

Our first clue that real networks may show manifestly nonrandom features also came 10 years ago from a map of the World Wide Web (WWW) (*8*), finding that the probability that a Web page has exactly k links (in other words, degree k) follows a power law distribution

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

a stunning departure from the Poisson distribution predicted by random network theory (*1*). Yet, it was not until we realized that Eq. 1 characterizes the network of actors linked by movies and scientific papers linked by citations (*9*) that we suspected that the scale-free property (*6*) might not be unique to the WWW. The main purpose of the 1999 *Science* paper was to report this unexpected similarity between networks of quite different nature and to show that two mechanisms, growth and preferential attachment, are the underlying causes (Fig. 1).

When we concluded in 1999 that we "expect that the scale invariant state […] is a generic property of many complex networks" (*7*), it was more of a prediction than a fact, because nature could have chosen as many different architectures as there are networks. Yet, probably the most surprising discovery of modern network theory is the universality of the network topology: Many real networks, from the cell to the Internet, independent of their age, function, and scope, converge to similar architectures. It is this universality that allowed researchers from different disciplines to embrace network theory as a common paradigm.

Today, the scale-free nature of networks of key scientific interest, from protein interactions to social networks and from the network of interlinked documents that make up the WWW to the interconnected hardware behind the Internet, has been established beyond doubt. The evidence comes not only from better maps and data sets but also from the agreement between empirical data and analytical models that predict the network structure (*10*, *11*). Yet, the early euphoria was not without negative side effects, prompting some researchers to label many systems scale-free, even when the evidence was scarce at best. However, the net result was to force us to better understand the factors that shape network structure. For ex-

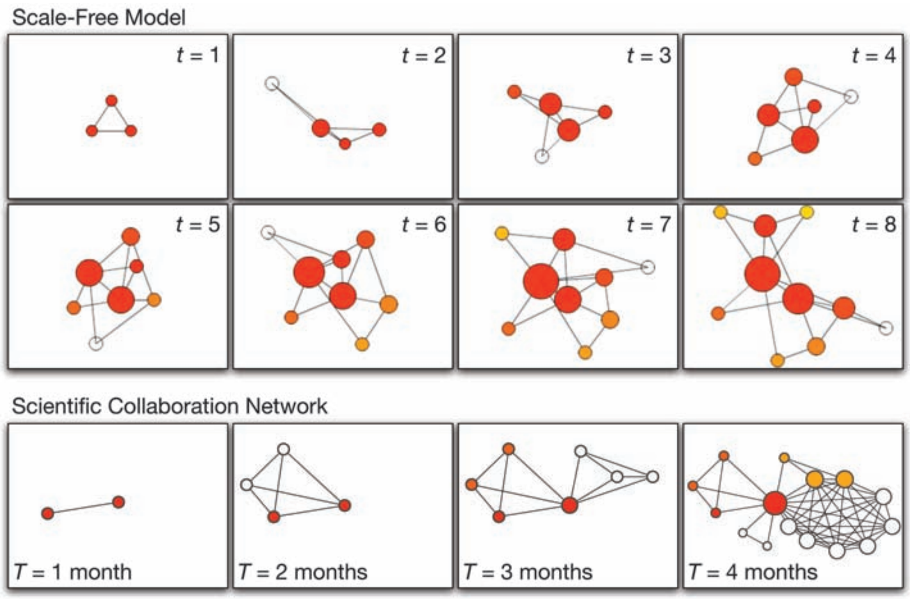

Fig. 1. The birth of a scale-free network. (**Top** and **Middle**) The simplest process that can produce a scale-free topology was introduced a decade ago in (*6*), and it is illustrated in the top two rows. Starting from three connected nodes (top left), in each image a new node (shown as an empty circle) is added to the network. When deciding where to link, new nodes prefer to attach to the more connected nodes, a process known as preferential attachment. Thanks to growth and preferential attachment, a rich-gets-richer process is observed, which means that the highly connected nodes acquire more links than those that are less connected, leading to the natural emergence of a few highly connected hubs. The node size, which was chosen to be proportional to the node's degree, illustrates the natural emergence of hubs as the largest nodes. The degree distribution of the resulting network follows the power law (Eq. 1) with exponent $\gamma = 3$. See also movies S1 to S3. (**Bottom**) Illustration of the growth process in the co-authorship network of physicists. Each node corresponds to an individual author, and two nodes are connected if they co-authored a paper together. The four images show the network's growth at 1-month time intervals, indicating how the network expands in time, leading to the emergence of a clear hub. Once again, the node size was chosen to be proportional to the node's degree. [Credit: D. Wang and G. Palla]

Center for Complex Network Research, Department of Physics, Biology, and Computer Science, Northeastern University, Boston, MA 02115, USA. Department of Medicine, Harvard Medical School and Center for Cancer Systems Biology, Dana Farber Cancer Institute, Boston, MA 02115, USA. E-mail: alb@neu.edu

ample, although the randomly bonded atoms in amorphous materials form a fascinating network, we now know that it does not display either the small-world (*12*) or the scale-free property, thanks to the chemical constraints the bonds must obey (*13*). Lastly, the topologies of several networks of considerable interest, like the neural-level map of a mammalian brain, remain to be elucidated, representing an area where we need both data and generative models (*14*).

A legacy of the scale-free property is the realization that the structure and the evolution of networks are inseparable (*6*). Indeed, traditional network models aimed to connect a fixed number of nodes with cleverly placed links. The scale-free property forced us to acknowledge that networks constantly change because of the arrival of nodes and links (Fig. 1). In other words, to explain a system's topology we first need to describe how it came into being.

The impact of network theory could have been limited if not for a series of findings that underlined the perils of ignoring network topology. Take, for example, the discovery of Romualdo Pastor-Satorras and Alessandro Vespignani that on a scale-free network the epidemic threshold converges to zero (*15*). It has long been known that only viruses whose spreading rate exceeds a critical threshold can survive in the population. Whereas the spreading rate captures the transmission dynamics, the threshold is determined by the topology of the network on which the virus spreads. Therefore, the vanishing threshold means that in scale-free networks even weakly virulent viruses can spread unopposed, a finding that affects all spreading processes, from AIDS to computer viruses. Similarly, the finding of Shlomo Havlin and collaborators (*16*) that in scale-free networks the overall network connectivity does not vanish under random node removal explained the exceptional robustness of real networks to random node failures (*17*). As a proof of the coherency of the emerging theory, both of these discoveries (*15*, *16*) were reduced to the same mathematical property, the diverging second moment of the degree distribution (Eq. 1), a unique feature of scale-free networks (*6*). Lately these features are of great interest, given the increasing concern about the vulnerability of real networks (such as power grids and the Internet) to attack and the realization that targeting hubs can be massively disruptive (*17*, *18*).

It is clear that no networks seen in nature or technology are completely random—that is, mechanisms beyond randomness shape their evolution. The universality of various topological characteristics, from degree distributions (*6*) to degree correlations (*19*–*21*), motifs (*22*), and communities (*23*–*25*), is used as a springboard to study diverse phenomena and to make predictions. With that, network theory has fundamentally reshaped our understanding of complexity. Indeed, although we continue to lack a universally agreed-on definition of complexity, the role of networks in this area is obvious: All systems perceived to be complex, from the cell to the Internet and from social to economic systems, consist of an extraordinarily large number of components that interact via intricate networks. To be sure, we were aware of these networks before. Yet, only recently have we acquired the data and tools to probe their topology, helping us realize that the underlying connectivity has such a strong impact on a system's behavior that no approach to complex systems can succeed unless it exploits the network topology.

In many ways, the demands of a future theory of complexity are obvious: We need to understand the behavior of the systems that we perceive as being complex. We need to be able to predict how the Internet responds to attacks and traffic jams or how the cell reacts to changes in its environment. To make progress in this direction, we need to tackle the next frontier, which is to understand the dynamics of the processes that take place on networks. The problem is that we have almost as many dynamical phenomena as there are complex systems. For example, biologists study reaction kinetics on metabolic networks; computer scientists monitor the flow of information on computer networks; and epidemiologists, sociologists, and economists explore the spread of viruses and ideas on social networks. Is there a chance that, despite their diversity, these dynamical processes share some common characteristics? I suspect that such commonalities do exist; we just have not yet found the framework to unveil their universality. If we do, combined with the universality of the network topology, we may soon have something that could form the foundation of a theory of complexity.

Can we keep the momentum and achieve this in the next decade or so? Perhaps—in my view the bottlenecks are mainly data driven. Indeed, the sudden emergence of large and reliable network maps drove the development of network theory during the past decade. If data of similar detail capturing the dynamics of processes taking place on networks were to emerge in the coming years, our imagination will be the only limitation to progress. If I dare to make a prediction for the next decade, it is this: Thanks to the proliferation of the many electronic devices that we use on a daily basis, from cell phones to Global Positioning Systems and the Internet, that capture everything from our communications to our whereabouts (*26*, *27*), the complex system that we are most likely to tackle first in a truly quantitative fashion may not be the cell or the Internet but rather society itself.

Today the understanding of networks is a common goal of an unprecedented array of traditional disciplines: Cell biologists use networks to make sense of signal transduction cascades and metabolism, to name a few applications in this area; computer scientists are mapping the Internet and the WWW; epidemiologists follow transmission networks through which viruses spread; and brain researchers are after the connectome, a neural-level connectivity map of the brain. Although many fads have come and gone in complexity, one thing is increasingly clear: Interconnectivity is so fundamental to the behavior of complex systems that networks are here to stay.

References and Notes

1. P. Erdős, A. Rényi, *Publ. Math. (Debrecen)* **6**, 290 (1959).
2. F. Karinthy, in *The Structure and Dynamics of Networks*, M. Newman, A.-L. Barabasi, D. Watts, Eds. (Princeton Univ. Press, Princeton, NJ, 2006).
3. S. Milgram, *Psychol. Today* **2**, 60 (1967).
4. I. Pool, M. Kochen, *Soc. Networks* **1**, 1 (1978).
5. D. J. Watts, S. H. Strogatz, *Nature* **393**, 440 (1998).
6. A.-L. Barabási, R. Albert, *Science* **286**, 509 (1999).
7. In a random network, the average node sets the scale of the network, which means that most nodes have about the same number of links as the average node. For networks that follow Eq. 1, for $\gamma < 3$ the second moment of the distribution diverges, which means that the average is not characteristic because the error bars characterizing our uncertainty about its value are infinite. These networks lack a characteristic scale; hence, they are called scale-free. Formally, networks whose degree distribution follows Eq. 1 are called scale-free networks.
8. R. Albert, H. Jeong, A.-L. Barabási, *Nature* **401**, 130 (1999).
9. S. Redner, *Eur. Phys. J. B* **4**, 131 (1998).
10. G. Caldarelli, *Scale-Free Networks* (Oxford Univ. Press, Oxford, 2007).
11. S. N. Dorogovtsev, J. F. F. Mendes, *Evolution of Networks: From Biological Networks to the Internet and WWW* (Oxford Univ. Press, Oxford, 2003).
12. The small-world property refers to the fact that in many networks the average node to node distance is rather small, of the order of log N, where N is the number of nodes in the network.
13. L. A. N. Amaral, A. Scala, M. Barthelemy, H. E. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11149 (2000).
14. E. T. Bullmore, O. Sporns, *Nat. Rev. Neurosci.* **10**, 186 (2009).
15. R. Pastor-Satorras, A. Vespignani, *Phys. Rev. Lett.* **86**, 3200 (2001).
16. R. Cohen, K. Reez, D. Ben-Avraham, S. Havlin, *Phys. Rev. Lett.* **85**, 4626 (2000).
17. R. Albert, H. Jeong, A.-L. Barabási, *Nature* **406**, 378 (2000).
18. A. E. Motter, *Phys. Rev. Lett.* **93**, 098701 (2004).
19. M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).
20. R. Pastor-Satorras, A. Vázquez, A. Vespignani, *Phys. Rev. Lett.* **87**, 258701 (2001).
21. S. Maslov, K. Sneppen, *Science* **296**, 910 (2002).
22. R. Milo *et al.*, *Science* **298**, 824 (2002).
23. M. E. J. Newman, *Phys. Today* **61**, 33 (2008).
24. G. Palla, I. Derényi, I. Farkas, T. Vicsek, *Nature* **435**, 814 (2005).
25. J. Reichardt, S. Bornholdt, *Phys. Rev. E* **74**, 016110 (2006).
26. A. Vespignani, *Science* **325**, 425 (2009).
27. M. C. González, C. A. Hidalgo, A.-L. Barabási, *Nature* **453**, 779 (2008).

Supporting Online Material
www.sciencemag.org/cgi/content/full/325/5939/412/DC1
Movies S1 to S3

10.1126/science.1173299