The structure and function of complex networks

M. E. J. Newman

Department of Physics, University of Michigan, Ann Arbor, MI 48109, U.S.A. and
Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, U.S.A.

Inspired by empirical studies of networked systems such as the Internet, social networks, and biological networks, researchers have in recent years developed a variety of techniques and models to help us understand or predict the behavior of these systems. Here we review developments in this field, including such concepts as the small-world effect, degree distributions, clustering, network correlations, random graph models, models of network growth and preferential attachment, and dynamical processes taking place on networks.

Contents

Acknowledgments

I. INTRODUCTION

A *network* is a set of items, which we will call *vertices* or sometimes nodes, with connections between them, called *edges* (Fig. 1). Systems taking the form of networks (also called "graphs" in much of the mathematical literature) abound in the world. Examples include the Internet, the World Wide Web, social networks of acquaintance or other connections between individuals, organizational networks and networks of business relations between companies, neural networks, metabolic networks, food webs, distribution networks such as blood vessels or postal delivery routes, networks of citations between papers, and many others (Fig. 2). This paper reviews recent (and some not-so-recent) work on the structure and function of networked systems such as these.

The study of networks, in the form of mathematical graph theory, is one of the fundamental pillars of discrete mathematics. Euler's celebrated 1735 solution of the Königsberg bridge problem is often cited as the first true proof in the theory of networks, and during the twentieth century graph theory has developed into a substantial body of knowledge.

Networks have also been studied extensively in the social sciences. Typical network studies in sociology involve the circulation of questionnaires, asking respondents to detail their interactions with others. One can then use the responses to reconstruct a network in which vertices represent individuals and edges the interactions between them. Typical social network studies address issues of centrality (which individuals are best connected to others or have most influence) and connectivity (whether and how individuals are connected to one another through the network).

Recent years however have witnessed a substantial new movement in network research, with the focus shifting away from the analysis of single small graphs and the properties of individual vertices or edges within such graphs to consideration of large-scale statistical properties of graphs. This new approach has been driven largely by the availability of computers and communication networks that allow us to gather and analyze data on a scale far larger than previously possible. Where studies used to look at networks of maybe tens or in extreme cases hundreds of vertices, it is not uncommon now to see networks with millions or even billions of vertices. This change of scale forces upon us a corresponding change in our analytic approach. Many of the questions that might previously have been asked in studies of small networks are simply not useful in much larger networks. A social network analyst might have asked, "Which vertex in this network would prove most crucial to the network's connectivity if it were removed?" But such a question has little meaning in most networks of a million vertices—no single vertex in such a network will have much effect at all when removed. On the other hand, one could reasonably ask a question like, "What percentage of vertices need to be removed to substantially affect network connectivity in some given way?" and this type of statistical question has real meaning even in a very large network.

However, there is another reason why our approach to the study of networks has changed in recent years, a reason whose importance should not be underestimated, although it often is. For networks of tens or hundreds of vertices, it is a relatively straightforward matter to draw a picture of the network with actual points and lines (Fig. 2) and to answer specific questions about network structure by examining this picture. This has been one of the primary methods of network analysts since the field began. The human eye is an analytic tool of remarkable power, and eyeballing pictures of networks is an excellent way to gain an understanding of their structure. With a network of a million or a billion vertices however, this approach is useless. One simply cannot draw a meaningful picture of a million vertices, even with modern 3D computer rendering tools, and therefore direct analysis by eye is hopeless. The recent development of statistical methods for quantifying large networks is to a large extent an attempt to find something to play the part played by the eye in the network analysis of the twentieth century. Statistical methods answer the question, "How can I tell what this network looks like, when I can't actually look at it?"

The body of theory that is the primary focus of this review aims to do three things. First, it aims to find statistical properties, such as path lengths and degree distributions, that characterize the structure and behavior of networked systems, and to suggest appropriate ways to measure these properties. Second, it aims to create models of networks that can help us to understand the meaning of these properties—how they came to be as they are, and how they interact with one another. Third, it aims to predict what the behavior of networked systems will be on the basis of measured structural properties and the local rules governing individual vertices. How for example will network structure affect traffic on the Internet, or the performance of a Web search engine, or the dynamics of social or biological systems? As we will see, the scientific community has, by drawing on ideas from a broad variety of disciplines, made an excellent start on the first two of these aims, the characterization and modeling of network structure. Studies of the effects of structure on system behavior on the other hand are still in their infancy. It remains to be seen what the crucial theoretical developments will be in this area.
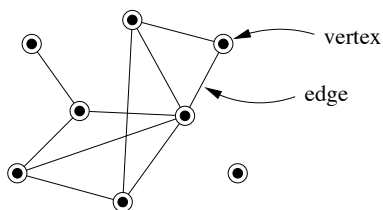
FIG. 1 A small example network with eight vertices and ten edges.
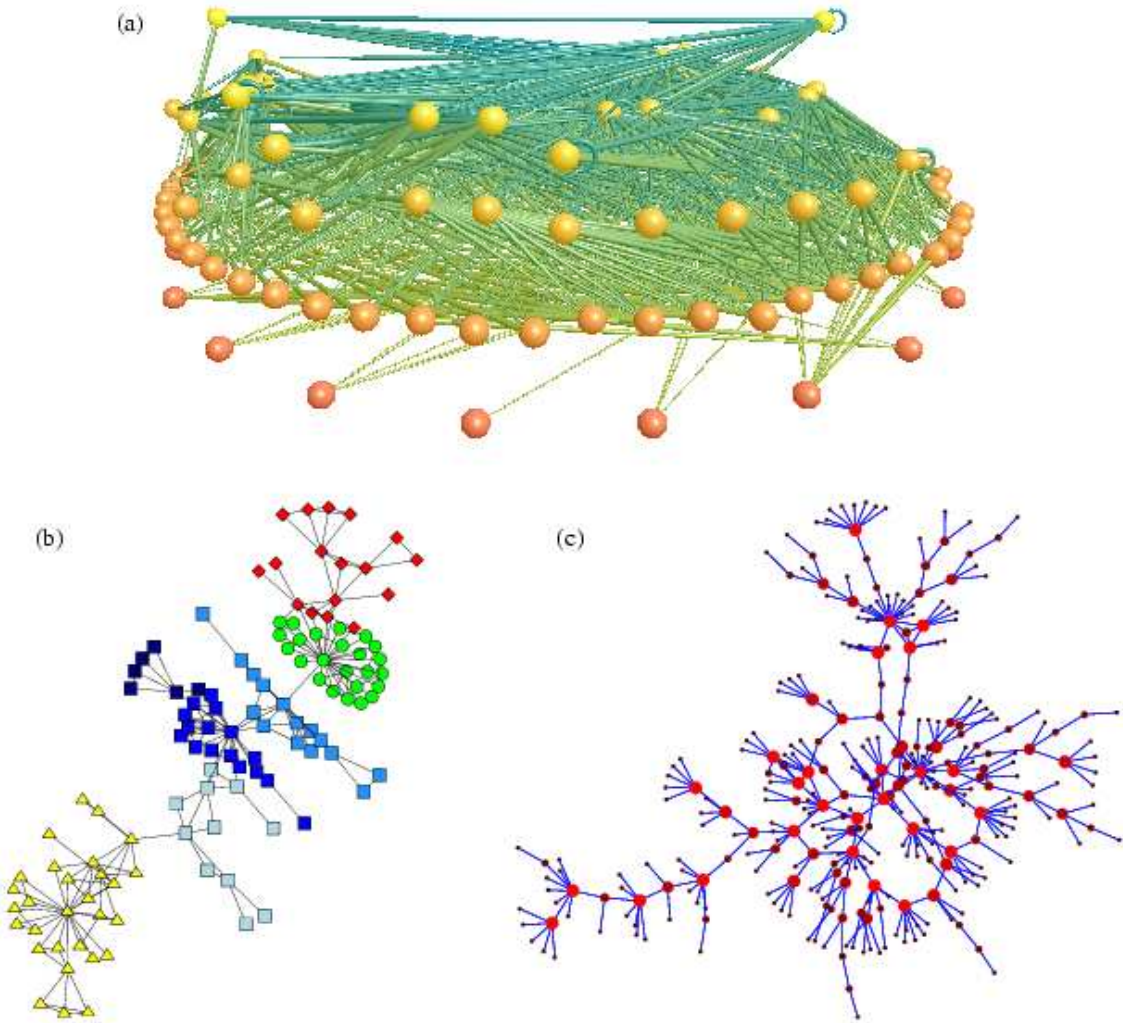
FIG. 2 Three examples of the kinds of networks that are the topic of this review. (a) A food web of predator-prey interactions between species in a freshwater lake [272]. Picture courtesy of Neo Martinez and Richard Williams. (b) The network of collaborations between scientists at a private research institution [171]. (c) A network of sexual contacts between individuals in the study by Potterat *et al.* [342].

A. Types of networks

A set of vertices joined by edges is only the simplest type of network; there are many ways in which networks may be more complex than this (Fig. 3). For instance, there may be more than one different type of vertex in a network, or more than one different type of edge. And vertices or edges may have a variety of properties, numerical or otherwise, associated with them. Taking the example of a social network of people, the vertices may represent men or women, people of different nationalities, locations, ages, incomes, or many other things. Edges may represent friendship, but they could also represent animosity, or professional acquaintance, or geographical proximity. They can carry weights, representing, say, how well two people know each other. They can also be directed, pointing in only one direction. Graphs composed of directed edges are themselves called directed

graphs or sometimes *digraphs*, for short. A graph representing telephone calls or email messages between individuals would be directed, since each message goes in only one direction. Directed graphs can be either cyclic, meaning they contain closed loops of edges, or acyclic meaning they do not. Some networks, such as food webs, are approximately but not perfectly acyclic.

One can also have *hyperedges*—edges that join more than two vertices together. Graphs containing such edges are called *hypergraphs*. Hyperedges could be used to indicate family ties in a social network for example—n individuals connected to each other by virtue of belonging to the same immediate family could be represented by an n-edge joining them. Graphs may also be naturally partitioned in various ways. We will see a number of examples in this review of *bipartite graphs*: graphs that contain vertices of two distinct types, with edges running only between unlike types. So-called *affiliation networks*

FIG. 3 Examples of various types of networks: (a) an undirected network with only a single type of vertex and a single type of edge; (b) a network with a number of discrete vertex and edge types; (c) a network with varying vertex and edge weights; (d) a directed network in which each edge has a direction.

in which people are joined together by common membership of groups take this form, the two types of vertices representing the people and the groups. Graphs may also evolve o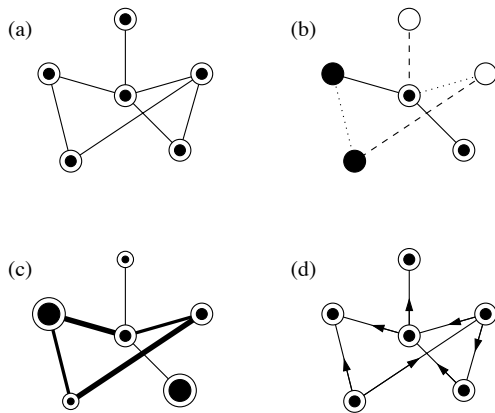ver time, with vertices or edges appearing or disappearing, or values defined on those vertices and edges changing. And there are many other levels of sophistication one can add. The study of networks is by no means a complete science yet, and many of the possibilities have yet to be explored in depth, but we will see examples of at least some of the variations described here in the work reviewed in this paper.

The jargon of the study of networks is unfortunately confused by differing usages among investigators from different fields. To avoid (or at least reduce) confusion, we give in Table II a short glossary of terms as they are used in this paper.

B. Other resources

A number of other reviews of this area have appeared recently, which the reader may wish to consult. Albert and Barabási [13] and Dorogovtsev and Mendes [120] have given extensive pedagogical reviews focusing on the physics literature. Both devote the larger part of their attention to the models of growing graphs that we describe in Sec. VII. Shorter reviews taking other viewpoints have been given by Newman [309] and Hayes [189, 190], who both concentrate on the so-called "small-world" models (see Sec. VI), and by Strogatz [387], who includes an interesting discussion of the behavior of dynamical systems on networks.

A number of books also make worthwhile reading. Dorogovtsev and Mendes [122] have expanded their above-mentioned review into a book, which again focuses on models of growing graphs. The edited volumes by Bornholdt and Schuster [70] and by Pastor-Satorras

and Rubi [330] both contain contributed essays on various topics by leading researchers. Detailed treatments of many of the topics covered in the present work can be found there. The book by Newman *et al.* [320] is a collection of previously published papers, and also contains some review material by the editors.

Three popular books on the subject of networks merit a mention. Albert-László Barabási's *Linked* [31] gives a personal account of recent developments in the study of networks, focusing particularly on Barabási's work on scale-free networks. Duncan Watts's *Six Degrees* [414] gives a sociologist's view, partly historical, of discoveries old and new. Mark Buchanan's *Nexus* [76] gives an entertaining portrait of the field from the point of view of a science journalist.

Farther afield, there are a variety of books on the study of networks in particular fields. Within graph theory the books by Harary [188] and by Bollobás [62] are widely cited and among social network theorists the books by Wasserman and Faust [409] and by Scott [363]. The book by Ahuja *et al.* [7] is a useful source for information on network algorithms.

C. Outline of the review

The outline of this paper is as follows. In Sec. II we describe empirical studies of the structure of networks, including social networks, information networks, technological networks and biological networks. In Sec. III we describe some of the common properties that are observed in many of these networks, how they are measured, and why they are believed to be important for the functioning of networked systems. Sections IV to VII form the heart of the review. They describe work on the mathematical modeling of networks, including random graph models and their generalizations, exponential random graphs, p* models and Markov graphs, the small-world model and its variations, and models of growing graphs including preferential attachment models and their many variations. In Sec. VIII we discuss the progress, such as it is, that has been made on the study of processes taking place on networks, including epidemic processes, network failure, models displaying phase transitions, and dynamical systems like random Boolean networks and cellular automata. In Sec. IX we give our conclusions and point to directions for future research.

II. NETWORKS IN THE REAL WORLD

In this section we look at what is known about the structure of networks of different types. Recent work on the mathematics of networks has been driven largely by observations of the properties of actual networks and attempts to model them, so network data are the obvious starting point for a review such as this. It also makes sense to examine simultaneously data from dif-

> *Vertex (pl. vertices):* The fundamental unit of a network, also called a site (physics), a node (computer science), or an actor (sociology).
>
> *Edge:* The line connecting two vertices. Also called a bond (physics), a link (computer science), or a tie (sociology).
>
> *Directed/undirected:* An edge is directed if it runs in only one direction (such as a one-way road between two points), and undirected if it runs in both directions. Directed edges, which are sometimes called *arcs*, can be thought of as sporting arrows indicating their orientation. A graph is directed if all of its edges are directed. An undirected graph can be represented by a directed one having two edges between each pair of connected vertices, one in each direction.
>
> *Degree:* The number of edges connected to a vertex. Note that the degree is not necessarily equal to the number of vertices adjacent to a vertex, since there may be more than one edge between any two vertices. In a few recent articles, the degree is referred to as the "connectivity" of a vertex, but we avoid this usage because the word connectivity already has another meaning in graph theory. A directed graph has both an in-degree and an out-degree for each vertex, which are the numbers of in-coming and out-going edges respectively.
>
> *Component:* The component to which a vertex belongs is that set of vertices that can be reached from it by paths running along edges of the graph. In a directed graph a vertex has both an in-component and an out-component, which are the sets of vertices from which the vertex can be reached and which can be reached from it.
>
> *Geodesic path:* A geodesic path is the shortest path through the network from one vertex to another. Note that there may be and often is more than one geodesic path between two vertices.
>
> *Diameter:* The diameter of a network is the length (in number of edges) of the longest geodesic path between any two vertices. A few authors have also used this term to mean the *average* geodesic distance in a graph, although strictly the two quantities are quite distinct.

TABLE I A short glossary of terms.

ferent kinds of networks. One of the principal thrusts of recent work in this area, inspired particularly by a groundbreaking 1998 paper by Watts and Strogatz [416], has been the comparative study of networks from different branches of science, with emphasis on properties that are common to many of them and the mathematical developments that mirror those properties. We here divide our summary into four loose categories of networks: social networks, information networks, technological networks and biological networks.

A. Social networks

A social network is a set of people or groups of people with some pattern of contacts or interactions between them [363, 409]. The patterns of friendships between individuals [296, 348], business relationships between companies [269, 286], and intermarriages between families [327] are all examples of networks that have been studied in the past.[1] Of the academic disciplines the so-cial sciences have the longest history of the substantial quantitative study of real-world networks [162, 363]. Of particular note among the early works on the subject are: Jacob Moreno's work in the 1920s and 30s on friendship patterns within small groups [296]; the so-called "southern women study" of Davis *et al.* [103], which focused on the social circles of women in an unnamed city in the American south in 1936; the study by Elton Mayo and colleagues of social networks of factory workers in the late 1930s in Chicago [357]; the mathematical models of Anatol Rapoport [346], who was one of the first theorists, perhaps *the* first, to stress the importance of the degree distribution in networks of all kinds, not just social networks; and the studies of friendship networks of school children by Rapoport and others [149, 348]. In more recent years, studies of business communities [167, 168, 269] and of patterns of sexual contacts [45, 218, 243, 266, 303, 342] have attracted particular attention.

Another important set of experiments are the famous

---

[1] Occasionally social networks of animals have been investigated also, such as dolphins [96], not to mention networks of fictional characters, such as the protagonists of Tolstoy's *Anna Karenina* [244] or Marvel Comics superheroes [10].

"small-world" experiments of Milgram [283, 393]. No actual networks were reconstructed in these experiments, but nonetheless they tell us about network structure. The experiments probed the distribution of path lengths in an acquaintance network by asking participants to pass a letter[2] to one of their first-name acquaintances in an attempt to get it to an assigned target individual. Most of the letters in the experiment were lost, but about a quarter reached the target and passed on average through the hands of only about six people in doing so. This experiment was the origin of the popular concept of the "six degrees of separation," although that phrase did not appear in Milgram's writing, being coined some decades later by Guare [183]. A brief but useful early review of Milgram's work and work stemming from it was given by Garfield [169].

Traditional social network studies often suffer from problems of inaccuracy, subjectivity, and small sample size. With the exception of a few ingenious indirect studies such as Milgram's, data collection is usually carried out by querying participants directly using questionnaires or interviews. Such methods are labor-intensive and therefore limit the size of the network that can be observed. Survey data are, moreover, influenced by subjective biases on the part of respondents; how one respondent defines a friend for example could be quite different from how another does. Although much effort is put into eliminating possible sources of inconsistency, it is generally accepted that there are large and essentially uncontrolled errors in most of these studies. A review of the issues has been given by Marsden [271].

Because of these problems many researchers have turned to other methods for probing social networks. One source of copious and relatively reliable data is collaboration networks. These are typically affiliation networks in which participants collaborate in groups of one kind or another, and links between pairs of individuals are established by common group membership. A classic, though rather frivolous, example of such a network is the collaboration network of film actors, which is thoroughly documented in the online Internet Movie Database.[3] In this network actors collaborate in films and two actors are considered connected if they have appeared in a film together. Statistical properties of this network have been analyzed by a number of authors [4, 20, 323, 416]. Other examples of networks of this type are networks of company directors, in which two directors are linked if they belong to the same board of directors [104, 105, 269], networks of coauthorship among academics, in which individuals are linked if they have coauthored one or more papers [36, 43, 68, 107, 182, 279, 292, 311, 312, 313], and coappearance networks in which individuals are linked by mention in the same context, particularly on Web

pages [3, 227] or in newspaper articles [99] (see Fig. 2b).

Another source of reliable data about personal connections between people is communication records of certain kinds. For example, one could construct a network in which each (directed) edge between two people represented a letter or package sent by mail from one to the other. No study of such a network has been published as far as we are aware, but some similar things have. Aiello *et al.* [8, 9] have analyzed a network of telephone calls made over the AT&T long-distance network on a single day. The vertices of this network represent telephone numbers and the directed edges calls from one number to another. Even for just a single day this graph is enormous, having about 50 million vertices, one of the largest graphs yet studied after the graph of the World Wide Web. Ebel *et al.* [136] have reconstructed the pattern of email communications between five thousand students at Kiel University from logs maintained by email servers. In this network the vertices represent email addresses and directed edges represent a message passing from one address to another. Email networks have also been studied by Newman *et al.* [321] and by Guimerà *et al.* [185], and similar networks have been constructed for an "instant messaging" system by Smith [371], and for an Internet community Web site by Holme *et al.* [196]. Dodds *et al.* [110] have carried out an email version of Milgram's small-world experiment in which participants were asked to forward an email message to one of their friends in an effort to get the message ultimately to some chosen target individual. Response rates for the experiment were quite low, but a few hundred completed chains of messages were recorded, enough to allow various statistical analyses.

B. Information networks

Our second network category is what we will call *information networks* (also sometimes called "knowledge networks"). The classic example of an information network is the network of citations between academic papers [138]. Most learned articles cite previous work by others on related topics. These citations form a network in which the vertices are articles and a directed edge from article A to article B indicates that A cites B. The structure of the citation network then reflects the structure of the information stored at its vertices, hence the term "information network," although certainly there are social aspects to the citation patterns of papers too [420].

Citation networks are acyclic (see Sec. I.A) because papers can only cite other papers that have already been written, not those that have yet to be written. Thus all edges in the network point backwards in time, making closed loops impossible, or at least extremely rare (see Fig. 4).

As an object of scientific study, citation networks have a great advantage in the copious and accurate data available for them. Quantitative study of publication patterns

---

[2] Actually a folder containing several documents.
[3] http://www.imdb.com/

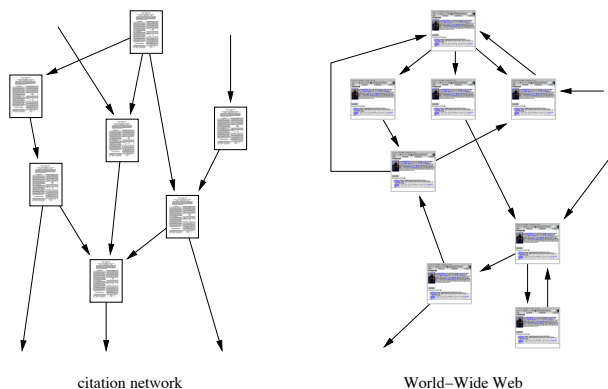citation network   World–Wide Web

FIG. 4 The two best studied information networks. Left: the citation network of academic papers in which the vertices are papers and the directed edges are citations of one paper by another. Since papers can only cite those that came before them (lower down in the figure) the graph is acyclic—it has no closed loops. Right: the World Wide Web, a network of text pages accessible over the Internet, in which the vertices are pages and the directed edges are hyperlinks. There are no constraints on the Web that forbid cycles and hence it is in general cyclic.

stretches back at least as far as Alfred Lotka's groundbreaking 1926 discovery of the so-called Law of Scientific Productivity, which states that the distribution of the numbers of papers written by individual scientists follows a power law. That is, the number of scientists who have written k papers falls off as $k^{-\alpha}$ for some constant α. (In fact, this result extends to the arts and humanities as well.) The first serious work on citation patterns was conducted in the 1960s as large citation databases became available through the work of Eugene Garfield and other pioneers in the field of bibliometrics. The network formed by citations was discussed in an early paper by Price [343], in which among other things, the author points out for the first time that both the in- and out-degree distributions of the network follow power laws, a far-reaching discovery which we discuss further in Sec. III.C. Many other studies of citation networks have been performed since then, using the ever better resources available in citation databases. Of particular note are the studies by Seglen [364] and Redner [351].[4]

Another very important example of an information network is the World Wide Web, which is a network of Web pages containing information, linked together by hyperlinks from one page to another [203]. The Web should not be confused with the Internet, which is a physical network of computers linked together by optical fibre and

other data connections.[5] Unlike a citation network, the World Wide Web is cyclic; there is no natural ordering of sites and no constraints that prevent the appearance of closed loops (Fig. 4). The Web has been very heavily studied since its first appearance in the early 1990s, with the studies by Albert *et al.* [14, 34], Kleinberg *et al.* [241], and Broder *et al.* [74] being particularly influential. The Web also appears to have power-law in- and out-degree distributions (Sec. III.C), as well as a variety of other interesting properties [2, 14, 74, 158, 241, 254].

One important point to notice about the Web is that our data about it come from "crawls" of the network, in which Web pages are found by following hyperlinks from other pages [74]. Our picture of the network structure of the World Wide Web is therefore necessarily biased. A page will only be found if another page points to it,[6] and in a crawl that covers only a part of the Web (as all crawls do at present) pages are more likely to be found the more other pages point to them [263]. This suggests for instance that our measurements of the fraction of pages with low in-degree might be an underestimate.[7] This behavior contrasts with that of a citation network. A paper can appear in the citation indices even if it has never been cited (and in fact a plurality of papers in the indices are never cited).

A few other examples of information networks have been studied to a lesser extent. Jaffe and Trajtenberg [207], for instance, have studied the network of citations between US patents, which is similar in some respects to citations between academic papers. A number of authors have looked at peer-to-peer networks [5, 6, 205], which are virtual networks of computers that allow sharing of files between computer users over local- or wide-area networks. The network of relations between word classes in a thesaurus has been studied by Knuth [244] and more recently by various other authors [234, 304, 384]. This network can be looked upon as an information network—users of a thesaurus "surf" the network from one word to another looking for the particular word that perfectly captures the idea they have in mind. However, it can also be looked at as a conceptual network representing the structure of the language, or possibly even the mental constructs used to represent the language. A number of other semantic word networks have also been investigated [119, 157, 369, 384].

Preference networks provide an example of a bipartite

---

[4] An interesting development in the study of citation patterns has been the arrival of automatic citation "crawlers" that construct citation networks from online papers. Examples include Citeseer (`http://citeseer.nj.nec.com/`), SPIRES (`http://www.slac.stanford.edu/spires/hep/`) and Citebase (`http://citebase.eprints.org/`).

[5] While the Web is primarily an information network, it, like citation networks, has social aspects to its structure also [3].

[6] This is not always strictly true. Some Web search engines allow the submission of pages by members of the public for inclusion in databases, and such pages need not be the target of links from any other pages. However, such pages also form a very small fraction of all Web pages, and certainly the biases discussed here remain very much present.

[7] The degree distribution for the Web shown in Fig. 6 falls off slightly at low values of the in-degree, which may perhaps reflect this bias.

information network. A preference network is a network with two kinds of vertices representing individuals and the objects of their preference, such as books or films, with an edge connecting each individual to the books or films they like. (Preference networks can also be weighted to indicate strength of likes or dislikes.) A widely studied example of a preference network is the *EachMovie* database of film preferences.[8] Networks of this kind form the basis for *collaborative filtering* algorithms and *recommender systems*, which are techniques for predicting new likes or dislikes based on comparison of individuals' preferences with those of others [176, 352, 367]. Collaborative filtering has found considerable commercial success for product recommendation and targeted advertising, particularly with online retailers. Preference networks can also be thought of as social networks, linking not only people to objects, but also people to other people with similar preferences. This approach has been adopted occasionally in the literature [227].

C. Technological networks

Our third class of networks is technological networks, man-made networks designed typically for distribution of some commodity or resource, such as electricity or information. The electric power grid is a good example. This is a network of high-voltage three-phase transmission lines that spans a country or a portion of a country (as opposed to the local low-voltage a.c. power delivery lines that span individual neighborhoods). Statistical studies of power grids have been made by, for example, Watts and Strogatz [412, 416] and Amaral *et al.* [20]. Other distribution networks that have been studied include the network of airline routes [20], and networks of roads [221], railways [262, 366] and pedestrian traffic [87]. River networks could be regarded as a naturally occurring form of distribution network (actually a collection network) [111, 270, 353, 356], as could the vascular networks discussed in Sec. II.D. The telephone network and delivery networks such as those used by the post-office or parcel delivery companies also fall into this general category and are presumably studied within the relevant corporations, if not yet by academic researchers. (We distinguish here between the physical telephone network of wires and cables and the network of who calls whom, discussed in Sec. II.A.) Electronic circuits [155] fall somewhere between distribution and communication networks.

Another very widely studied technological network is the Internet, i.e., the network of physical connections between computers. Since there is a large and ever-changing number of computers on the Internet, the structure of the network is usually examined at a coarse-grained level, either the level of routers, special-purpose computers on the network that control the movement of data, or "autonomous systems," which are groups of computers within which networking is handled locally, but between which data flows over the public Internet. The computers at a single company or university would probably form a single autonomous system—autonomous systems often correspond roughly with domain names.

In fact, the network of physical connections on the Internet is not easy to discover since the infrastructure is maintained by many separate organizations. Typically therefore, researchers reconstruct the network by reasoning from large samples of point-to-point data routes. So-called "traceroute" programs can report the sequence of network nodes that a data packet passes through when traveling between two points and if we assume an edge in the network between any two consecutive nodes along such a path then a sufficiently large sample of paths will give us a fairly complete picture of the entire network. There may however be some edges that never get sampled, so the reconstruction is typically a good, but not perfect, representation of the true physical structure of the Internet. Studies of Internet structure have been carried out by, among others, Faloutsos *et al.* [148], Broida and Claffy [75] and Chen *et al.* [86].

D. Biological networks

A number of biological systems can be usefully represented as networks. Perhaps the classic example of a biological network is the network of metabolic pathways, which is a representation of metabolic substrates and products with directed edges joining them if a known metabolic reaction exists that acts on a given substrate and produces a given product. Most of us will probably have seen at some point the giant maps of metabolic pathways that many molecular biologists pin to their walls.[9] Studies of the statistical properties of metabolic networks have been performed by, for example, Jeong *et al.* [214, 340], Fell and Wagner [153, 405], and Stelling *et al.* [383]. A separate network is the network of mechanistic physical interactions between proteins (as opposed to chemical reactions among metabolites), which is usually referred to as a protein interaction network. Interaction networks have been studied by a number of authors [206, 212, 274, 376, 394].

Another important class of biological network is the genetic regulatory network. The expression of a gene, i.e., the production by transcription and translation of the protein for which the gene codes, can be controlled by the presence of other proteins, both activators and

---

[8] http://research.compaq.com/SRC/eachmovie/

[9] The standard chart of the metabolic network is somewhat misleading. For reasons of clarity and aesthetics, many metabolites appear in more than one place on the chart, so that some pairs of vertices are actually the same vertex.

inhibitors, so that the genome itself forms a switching network with vertices representing the proteins and directed edges representing dependence of protein production on the proteins at other vertices. The statistical structure of regulatory networks has been studied recently by various authors [152, 184, 368]. Genetic regulatory networks were in fact one of the first networked dynamical systems for which large-scale modeling attempts were made. The early work on random Boolean nets by Kauffman [224, 225, 226] is a classic in this field, and anticipated recent developments by several decades.

Another much studied example of a biological network is the food web, in which the vertices represent species in an ecosystem and a directed edge from species A to species B indicates that A preys on B [91, 339]—see Fig. 2a. (Sometimes the relationship is drawn the other way around, because ecologists tend to think in terms of energy or carbon flows through food webs; a predator-prey interaction is thus drawn as an arrow pointing from prey to predator, indicating energy flow from prey to predator when the prey is eaten.) Construction of complete food webs is a laborious business, but a number of quite extensive data sets have become available in recent years [27, 177, 204, 272]. Statistical studies of the topologies of food webs have been carried out by Solé and Montoya [290, 375], Camacho *et al.* [82] and Dunne *et al.* [132, 133, 423], among others. A particularly thorough study of webs of plants and herbivores has been conducted by Jordano *et al.* [219], which includes statistics for no less than 53 different networks.

Neural networks are another class of biological networks of considerable importance. Measuring the topology of real neural networks is extremely difficult, but has been done successfully in a few cases. The best known example is the reconstruction of the 282-neuron neural network of the nematode *C. Elegans* by White *et al.* [421]. The network structure of the brain at larger scales than individual neurons—functional areas and pathways—has been investigated by Sporns *et al.* [379, 380].

Blood vessels and the equivalent vascular networks in plants form the foundation for one of the most successful theoretical models of the effects of network structure on the behavior of a networked system, the theory of biological allometry [29, 417, 418], although we are not aware of any quantitative studies of their statistical structure.

Finally we mention two examples of networks from the physical sciences, the network of free energy minima and saddle points in glasses [130] and the network of conformations of polymers and the transitions between them [361], both of which appear to have some interesting structural properties.

III. PROPERTIES OF NETWORKS

Perhaps the simplest useful model of a network is the random graph, first studied by Rapoport [346, 347, 378] and by Erdős and Rényi [141, 142, 143], which we describe in Sec. IV.A. In this model, undirected edges are placed at random between a fixed number n of vertices to create a network in which each of the $\frac{1}{2}n(n-1)$ possible edges is independently present with some probability p, and the number of edges connected to each vertex—the degree of the vertex—is distributed according to a binomial distribution, or a Poisson distribution in the limit of large n. The random graph has been well studied by mathematicians [63, 211, 223] and many results, both approximate and exact, have been proved rigorously. Most of the interesting features of real-world networks that have attracted the attention of researchers in the last few years however concern the ways in which networks are *not* like random graphs. Real networks are non-random in some revealing ways that suggest both possible mechanisms that could be guiding network formation, and possible ways in which we could exploit network structure to achieve certain aims. In this section we describe some features that appear to be common to networks of many different types.

A. The small-world effect

In Sec. II.A we described the famous experiments carried out by Stanley Milgram in the 1960s, in which letters passed from person to person were able to reach a designated target individual in only a small number of steps—around six in the published cases. This result is one of the first direct demonstrations of the *small-world effect*, the fact that most pairs of vertices in most networks seem to be connected by a short path through the network.

The existence of the small-world effect had been speculated upon before Milgram's work, notably in a remarkable 1929 short story by the Hungarian writer Frigyes Karinthy [222], and more rigorously in the mathematical work of Pool and Kochen [341] which, although published after Milgram's studies, was in circulation in preprint form for a decade before Milgram took up the problem. Nowadays, the small-world effect has been studied and verified directly in a large number of different networks.

Consider an undirected network, and let us define ℓ to be the mean geodesic (i.e., shortest) distance between vertex pairs in a network:

$$\ell = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \geq j} d_{ij}, \qquad (1)$$

where d_{ij} is the geodesic distance from vertex i to vertex j. Notice that we have included the distance from each vertex to itself (which is zero) in this average. This is mathematically convenient for a number of reasons, but not all authors do it. In any case, its inclusion simply multiplies ℓ by $(n-1)/(n+1)$ and hence gives a correction of order n^{-1}, which is often negligible for practical purposes.

The quantity ℓ can be measured for a network of n vertices and m edges in time $O(mn)$ using simple breadth-

	network	type	n	m	z	ℓ	α	$C^{(1)}$	$C^{(2)}$	r	Ref(s).
social	film actors	undirected	449 913	25 516 482	113.43	3.48	2.3	0.20	0.78	0.208	20, 416
	company directors	undirected	7 673	55 392	14.44	4.60	–	0.59	0.88	0.276	105, 323
	math coauthorship	undirected	253 339	496 489	3.92	7.57	–	0.15	0.34	0.120	107, 182
	physics coauthorship	undirected	52 909	245 300	9.27	6.19	–	0.45	0.56	0.363	311, 313
	biology coauthorship	undirected	1 520 251	11 803 064	15.53	4.92	–	0.088	0.60	0.127	311, 313
	telephone call graph	undirected	47 000 000	80 000 000	3.16		2.1				8, 9
	email messages	directed	59 912	86 300	1.44	4.95	1.5/2.0		0.16		136
	email address books	directed	16 881	57 029	3.38	5.22	–	0.17	0.13	0.092	321
	student relationships	undirected	573	477	1.66	16.01	–	0.005	0.001	−0.029	45
	sexual contacts	undirected	2 810				3.2				265, 266
information	WWW nd.edu	directed	269 504	1 497 135	5.55	11.27	2.1/2.4	0.11	0.29	−0.067	14, 34
	WWW Altavista	directed	203 549 046	2 130 000 000	10.46	16.18	2.1/2.7				74
	citation network	directed	783 339	6 716 198	8.57		3.0/–				351
	Roget's Thesaurus	directed	1 022	5 103	4.99	4.87	–	0.13	0.15	0.157	244
	word co-occurrence	undirected	460 902	17 000 000	70.13		2.7		0.44		119, 157
technological	Internet	undirected	10 697	31 992	5.98	3.31	2.5	0.035	0.39	−0.189	86, 148
	power grid	undirected	4 941	6 594	2.67	18.99	–	0.10	0.080	−0.003	416
	train routes	undirected	587	19 603	66.79	2.16	–		0.69	−0.033	366
	software packages	directed	1 439	1 723	1.20	2.42	1.6/1.4	0.070	0.082	−0.016	318
	software classes	directed	1 377	2 213	1.61	1.51	–	0.033	0.012	−0.119	395
	electronic circuits	undirected	24 097	53 248	4.34	11.05	3.0	0.010	0.030	−0.154	155
	peer-to-peer network	undirected	880	1 296	1.47	4.28	2.1	0.012	0.011	−0.366	6, 354
biological	metabolic network	undirected	765	3 686	9.64	2.56	2.2	0.090	0.67	−0.240	214
	protein interactions	undirected	2 115	2 240	2.12	6.80	2.4	0.072	0.071	−0.156	212
	marine food web	directed	135	598	4.43	2.05	–	0.16	0.23	−0.263	204
	freshwater food web	directed	92	997	10.84	1.90	–	0.20	0.087	−0.326	272
	neural network	directed	307	2 359	7.68	3.97	–	0.18	0.28	−0.226	416, 421

TABLE II Basic statistics for a number of published networks. The properties measured are: type of graph, directed or undirected; total number of vertices n; total number of edges m; mean degree z; mean vertex–vertex distance ℓ; exponent α of degree distribution if the distribution follows a power law (or "–" if not; in/out-degree exponents are given for directed graphs); clustering coefficient $C^{(1)}$ from Eq. (3); clustering coefficient $C^{(2)}$ from Eq. (6); and degree correlation coefficient r, Sec. III.F. The last column gives the citation(s) for the network in the bibliography. Blank entries indicate unavailable data.

first search [7], also called a "burning algorithm" in the physics literature. In Table II, we show values of ℓ taken from the literature for a variety of different networks. As the table shows, the values are in all cases quite small—much smaller than the number n of vertices, for instance.

The definition (1) of ℓ is problematic in networks that have more than one component. In such cases, there exist vertex pairs that have no connecting path. Conventionally one assigns infinite geodesic distance to such pairs, but then the value of ℓ also becomes infinite. To avoid this problem one usually defines ℓ on such networks to be the mean geodesic distance between all pairs that have a connecting path. Pairs that fall in two different components are excluded from the average. The figures in Table II were all calculated in this way. An alternative and perhaps more satisfactory approach is to define ℓ to be the "harmonic mean" geodesic distance between all pairs, i.e., the reciprocal of the average of the reciprocals:

$$\ell^{-1} = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \geq j} d_{ij}^{-1}. \qquad (2)$$

Infinite values of d_{ij} then contribute nothing to the sum. This approach has been adopted only occasionally in network calculations [260], but perhaps should be used more often.

The small-world effect has obvious implications for the dynamics of processes taking place on networks. For example, if one considers the spread of information, or indeed anything else, across a network, the small-world effect implies that that spread will be fast on most real-world networks. If it takes only six steps for a rumor to spread from any person to any other, for instance, then the rumor will spread much faster than if it takes a hundred steps, or a million. This affects the number of "hops" a packet must make to get from one computer to another on the Internet, the number of legs of a journey for an air or train traveler, the time it takes for a disease to spread throughout a population, and so forth. The small-world effect also underlies some well-known parlor games, particularly the calculation of Erdős numbers [107] and Bacon numbers.[10]

On the other hand, the small-world effect is also mathematically obvious. If the number of vertices within a distance r of a typical central vertex grows exponentially with r—and this is true of many networks, including the random graph (Sec. IV.A)—then the value of ℓ will increase as $\log n$. In recent years the term "small-world effect" has thus taken on a more precise meaning: networks are said to show the small-world effect if the value of ℓ scales logarithmically or slower with network size for fixed mean degree. Logarithmic scaling can be proved for a variety of network models [61, 63, 88, 127, 164]
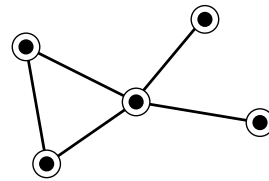
---

[10] http://www.cs.virginia.edu/oracle/

FIG. 5 Illustration of the definition of the clustering coefficient C, Eq. (3). This network has one triangle and eight connected triples, and therefore has a clustering coefficient of $3 \times 1/8 = \frac{3}{8}$. The individual vertices have local clustering coefficients, Eq. (5), of 1, 1, $\frac{1}{6}$, 0 and 0, for a mean value, Eq. (6), of $C = \frac{13}{30}$.

and has also been observed in various real-world networks [13, 312, 313]. Some networks have mean vertex–vertex distances that increase slower than $\log n$. Bollobás and Riordan [64] have shown that networks with power-law degree distributions (Sec. III.C) have values of ℓ that increase no faster than $\log n / \log \log n$ (see also Ref. [164]), and Cohen and Havlin [95] have given arguments that suggest that the actual variation may be slower even than this.

B. Transitivity or clustering

A clear deviation from the behavior of the random graph can be seen in the property of network transitivity, sometimes also called clustering, although the latter term also has another meaning in the study of networks (see Sec. III.G) and so can be confusing. In many networks it is found that if vertex A is connected to vertex B and vertex B to vertex C, then there is a heightened probability that vertex A will also be connected to vertex C. In the language of social networks, the friend of your friend is likely also to be your friend. In terms of network topology, transitivity means the presence of a heightened number of triangles in the network—sets of three vertices each of which is connected to each of the others. It can be quantified by defining a clustering coefficient C thus:

$$C = \frac{3 \times \text{ number of triangles in the network}}{\text{number of connected triples of vertices}}, \qquad (3)$$

where a "connected triple" means a single vertex with edges running to an unordered pair of others (see Fig. 5).

In effect, C measures the fraction of triples that have their third edge filled in to complete the triangle. The factor of three in the numerator accounts for the fact that each triangle contributes to three triples and ensures that C lies in the range $0 \leq C \leq 1$. In simple terms, C is the mean probability that two vertices that are network neighbors of the same other vertex will themselves be neighbors. It can also be written in the form

$$C = \frac{6 \times \text{ number of triangles in the network}}{\text{number of paths of length two}}, \qquad (4)$$

where a path of length two refers to a directed path starting from a specified vertex. This definition shows that C is also the mean probability that the friend of your friend is also your friend.

The definition of C given here has been widely used in the sociology literature, where it is referred to as the "fraction of transitive triples."[11] In the mathematical and physical literature it seems to have been first discussed by Barrat and Weigt [40].

An alternative definition of the clustering coefficient, also widely used, has been given by Watts and Strogatz [416], who proposed defining a local value

$$C_i = \frac{\text{number of triangles connected to vertex } i}{\text{number of triples centered on vertex } i}. \quad (5)$$

For vertices with degree 0 or 1, for which both numerator and denominator are zero, we put $C_i = 0$. Then the clustering coefficient for the whole network is the average

$$C = \frac{1}{n} \sum_i C_i. \quad (6)$$

This definition effectively reverses the order of the operations of taking the ratio of triangles to triples and of averaging over vertices—one here calculates the mean of the ratio, rather than the ratio of the means. It tends to weight the contributions of low-degree vertices more heavily, because such vertices have a small denominator in Eq. (5) and hence can give quite different results from Eq. (3). In Table II we give both measures for a number of networks (denoted $C^{(1)}$ and $C^{(2)}$ in the table). Normally our first definition (3) is easier to calculate analytically, but (6) is easily calculated on a computer and has found wide use in numerical studies and data analysis. It is important when reading (or writing) literature in this area to be clear about which definition of the clustering coefficient is in use. The difference between the two is illustrated in Fig. 5.

The local clustering C_i above has been used quite widely in its own right in the sociological literature, where it is referred to as the "network density" [363]. Its dependence on the degree k_i of the central vertex i has been studied by Dorogovtsev *et al.* [113] and Szabó *et al.* [389]; both groups found that C_i falls off with k_i approximately as k_i^{-1} for certain models of scale-free networks (Sec. III.C.1). Similar behavior has also been observed empirically in real-world networks [349, 350, 397].

In general, regardless of which definition of the clustering coefficient is used, the values tend to be considerably higher than for a random graph with a similar number of vertices and edges. Indeed, it is suspected

that for many types of networks the probability that the friend of your friend is also your friend should tend to a non-zero limit as the network becomes large, so that $C = \mathrm{O}(1)$ as $n \to \infty$.[12] On the random graph, by contrast, $C = \mathrm{O}(n^{-1})$ for large n (either definition of C) and hence the real-world and random graph values can be expected to differ by a factor of order n. This point is discussed further in Sec. IV.A.

The clustering coefficient measures the density of triangles in a network. An obvious generalization is to ask about the density of longer loops also: loops of length four and above. A number of authors have looked at such higher order clustering coefficients [54, 79, 165, 172, 317], although there is so far no clean theory, similar to a cumulant expansion, that separates the independent contributions of the various orders from one another. If more than one edge is permitted between a pair of vertices, then there is also a lower order clustering coefficient that describes the density of loops of length two. This coefficient is particularly important in directed graphs where the two edges in question can point in opposite directions. The probability that two vertices in a directed network point to each other is called the *reciprocity* and is often measured in directed social networks [363, 409]. It has been examined occasionally in other contexts too, such as the World Wide Web [3, 137] and email networks [321].

C. Degree distributions

Recall that the degree of a vertex in a network is the number of edges incident on (i.e., connected to) that vertex. We define p_k to be the fraction of vertices in the network that have degree k. Equivalently, p_k is the probability that a vertex chosen uniformly at random has degree k. A plot of p_k for any given network can be formed by making a histogram of the degrees of vertices. This histogram is the degree distribution for the network. In a random graph of the type studied by Erdős and Rényi [141, 142, 143], each edge is present or absent with equal probability, and hence the degree distribution is, as mentioned earlier, binomial, or Poisson in the limit of large graph size. Real-world networks are mostly found to be very unlike the random graph in their degree distributions. Far from having a Poisson distribution, the degrees of the vertices in most networks are highly right-skewed, meaning that their distribution has a long right tail of values that are far above the mean.

Measuring this tail is somewhat tricky. Although in theory one just has to construct a histogram of the degrees, in practice one rarely has enough measurements to get good statistics in the tail, and direct histograms are

---

[11] For example, the standard network analysis program UCInet includes a function to calculate this quantity for any network.

[12] An exception is scale-free networks with $C_i \sim k_i^{-1}$, as described above. For such networks Eq. (3) tends to zero as $n \to \infty$, although Eq. (6) is still finite.

thus usually rather noisy (see the histograms in Refs. [74, 148] and [343] for example). There are two accepted ways to get around this problem. One is to constructed a histogram in which the bin sizes increase exponentially with degree. For example the first few bins might cover degree ranges 1, 2–3, 4–7, 8–15, and so on. The number of samples in each bin is then divided by the width of the bin to normalize the measurement. This method of constructing a histogram is often used when the histogram is to be plotted with a logarithmic degree scale, so that the widths of the bins will appear even. Because the bins get wider as we get out into the tail, the problems with statistics are reduced, although they are still present to some extent as long as p_k falls off faster than k^{-1}, which it must if the distribution is to be integrable.

An alternative way of presenting degree data is to make a plot of the cumulative distribution function

$$P_k = \sum_{k'=k}^{\infty} p_{k'}, \qquad (7)$$

which is the probability that the degree is greater than or equal to k. Such a plot has the advantage that all the original data are represented. When we make a conventional histogram by binning, any differences between the values of data points that fall in the same bin are lost. The cumulative distribution function does not suffer from this problem. The cumulative distribution also reduces the noise in the tail. On the downside, the plot doesn't give a direct visualization of the degree distribution itself, and adjacent points on the plot are not statistically independent, making correct fits to the data tricky.

In Fig. 6 we show cumulative distributions of degree for a number of the networks described in Sec. II. As the figure shows, the distributions are indeed all right-skewed. Many of them follow power laws in their tails: $p_k \sim k^{-\alpha}$ for some constant exponent α. Note that such power-law distributions show up as power laws in the cumulative distributions also, but with exponent $\alpha - 1$ rather than α:

$$P_k \sim \sum_{k'=k}^{\infty} k'^{-\alpha} \sim k^{-(\alpha-1)}. \qquad (8)$$

Some of the other distributions have exponential tails: $p_k \sim e^{-k/\kappa}$. These also give exponentials in the cumulative distribution, but with the *same* exponent:

$$P_k = \sum_{k'=k}^{\infty} p_k \sim \sum_{k'=k}^{\infty} e^{-k'/\kappa} \sim e^{-k/\kappa}. \qquad (9)$$

This makes power-law and exponential distributions particularly easy to spot experimentally, by plotting the corresponding cumulative distributions on logarithmic scales (for power laws) or semi-logarithmic scales (for exponentials).

For other types of networks degree distributions can be more complicated. For bipartite graphs, for instance

(Sec. I.A), there are two degree distributions, one for each type of vertex. For directed graphs each vertex has both an in-degree and an out-degree, and the degree distribution therefore becomes a function p_{jk} of two variables, representing the fraction of vertices that simultaneously have in-degree j and out-degree k. In empirical studies of directed graphs like the Web, researchers have usually given only the individual distributions of in- and out-degree [14, 34, 74], i.e., the distributions derived by summing p_{jk} over one or other of its indices. This however discards much of the information present in the joint distribution. It has been found that in- and out-degrees are quite strongly correlated in some networks [321], which suggests that there is more to be gleaned from the joint distribution than is normally appreciated.

1. Scale-free networks

Networks with power-law degree distributions have been the focus of a great deal of attention in the literature [13, 120, 387]. They are sometimes referred to as *scale-free networks* [32], although it is only their degree distributions that are scale-free;[13] one can and usually does have scales present in other network properties. The earliest published example of a scale-free network is probably Price's network of citations between scientific papers [343] (see Sec. II.B). He quoted a value of $\alpha = 2.5$ to 3 for the exponent of his network. In a later paper he quoted a more accurate figure of $\alpha = 3.04$ [344]. He also found a power-law distribution for the out-degree of the network (number of bibliography entries in each paper), although later work has called this into question [396]. More recently, power-law degree distributions have been observed in a host of other networks, including notably other citation networks [351, 364], the World Wide Web [14, 34, 74], the Internet [86, 148, 401], metabolic networks [212, 214], telephone call graphs [8, 9], and the network of human sexual contacts [218, 266]. The degree distributions of some of these networks are shown in Fig. 6.

Other common functional forms for the degree distribution are exponentials, such as those seen in the power grid [20] and railway networks [366], and power laws with exponential cutoffs, such as those seen in the network of movie actors [20] and some collaboration networks [313]. Note also that while a particular form may be seen in the degree distribution for the network as a whole, specific subnetworks within the network can have other forms. The World Wide Web, for instance, shows a power-law

---

[13] The term "scale-free" refers to any functional form $f(x)$ that remains unchanged to within a multiplicative factor under a rescaling of the independent variable x. In effect this means power-law forms, since these are the only solutions to $f(ax) = bf(x)$, and hence "power-law" and "scale-free" are, for our purposes, synonymous.
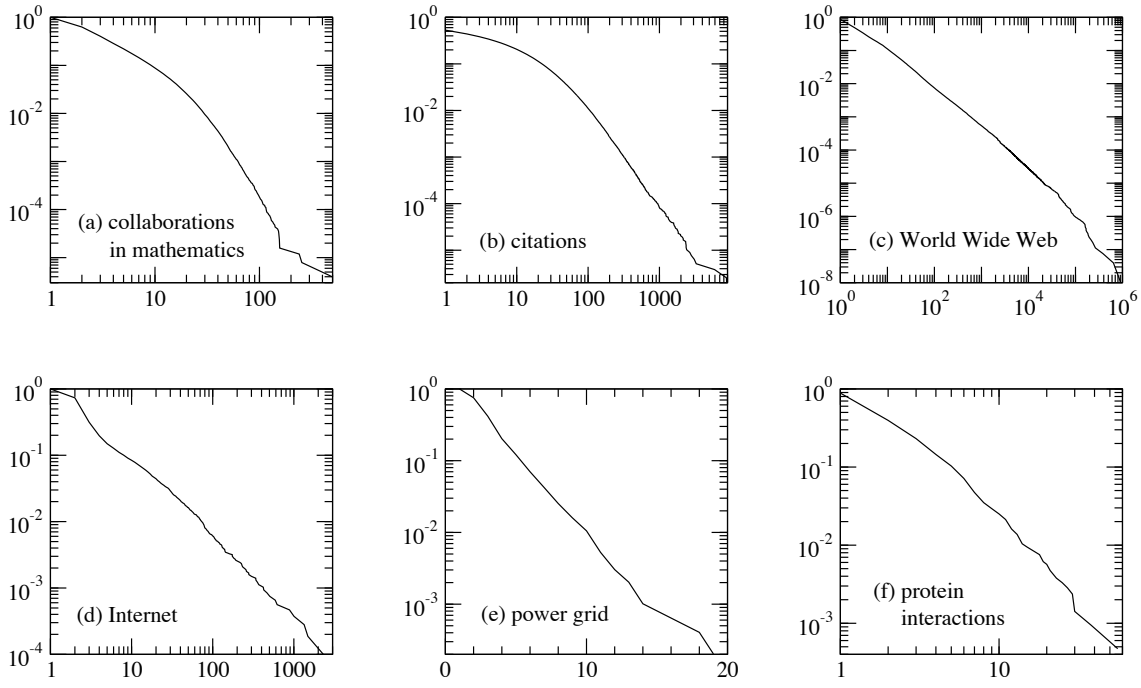
FIG. 6 Cumulative degree distributions for six different networks. The horizontal axis for each panel is vertex degree k (or in-degree for the citation and Web networks, which are directed) and the vertical axis is the cumulative probability distribution of degrees, i.e., the fraction of vertices that have degree greater than or equal to k. The networks shown are: (a) the collaboration network of mathematicians [182]; (b) citations between 1981 and 1997 to all papers cataloged by the Institute for Scientific Information [351]; (c) a 300 million vertex subset of the World Wide Web, *circa* 1999 [74]; (d) the Internet at the level of autonomous systems, April 1999 [86]; (e) the power grid of the western United States [416]; (f) the interaction network of proteins in the metabolism of the yeast *S. Cerevisiae* [212]. Of these networks, three of them, (c), (d) and (f), appear to have power-law degree distributions, as indicated by their approximately straight-line forms on the doubly logarithmic scales, and one (b) has a power-law tail but deviates markedly from power-law behavior for small degree. Network (e) has an exponential degree distribution (note the log-linear scales used in this panel) and network (a) appears to have a truncated power-law degree distribution of some type, or possibly two separate power-law regimes with different exponents.

degree distribution overall but unimodal distributions within domains [338].

2. Maximum degree

The maximum degree k_{max} of a vertex in a network will in general depend on the size of the network. For some calculations on networks the value of this maximum degree matters (see, for example, Sec. VIII.C.2). In work on scale-free networks, Aiello *et al.* [8] assumed that the maximum degree was approximately the value above which there is less than one vertex of that degree in the graph on average, i.e., the point where $np_k = 1$. This means, for instance, that $k_{\mathrm{max}} \sim n^{1/\alpha}$ for the power-law degree distribution $p_k \sim k^{-\alpha}$. This assumption however can give misleading results; in many cases there will be vertices in the network with significantly higher degree than this, as discussed by Adamic *et al.* [6].

Given a particular degree distribution (and assuming all degrees to be sampled independently from it, which may not be true for networks in the real world), the probability of there being exactly m vertices of degree k and

no vertices of higher degree is $\binom{n}{m} p_k^m (1 - P_k)^{n-m}$, where P_k is the cumulative probability distribution, Eq. (7). Hence the probability h_k that the highest degree on the graph is k is

$$h_k = \sum_{m=1}^{n} \binom{n}{m} p_k^m (1 - P_k)^{n-m}$$
$$= (p_k + 1 - P_k)^n - (1 - P_k)^n, \qquad (10)$$

and the expected value of the highest degree is $k_{\mathrm{max}} = \sum_k k h_k$.

For both small and large values of k, h_k tends to zero, and the sum over k is dominated by the terms close to the maximum. Thus, in most cases, a good approximation to the expected value of the maximum degree is given by the modal value. Differentiating and observing that $\mathrm{d}P_k/\mathrm{d}k = p_k$, we find that the maximum of h_k occurs when

$$\left(\frac{\mathrm{d}p_k}{\mathrm{d}k} - p_k\right)(p_k + 1 - P_k)^{n-1} + p_k(1 - P_k)^{n-1} = 0, \quad (11)$$

or k_{max} is a solution of

$$\frac{\mathrm{d}p_k}{\mathrm{d}k} \simeq -np_k^2, \qquad (12)$$

where we have made the (fairly safe) assumption that p_k is sufficiently small for $k \gtrsim k_{max}$ that $np_k \ll 1$ and $P_k \ll 1$.

For example, if $p_k \sim k^{-\alpha}$ in its tail, then we find that

$$k_{max} \sim n^{1/(\alpha-1)}. \qquad (13)$$

As shown by Cohen *et al.* [93], a simple rule of thumb that leads to the same result is that the maximum degree is roughly the value of k that solves $nP_k = 1$. Note however that, as shown by Dorogovtsev and Samukhin [129], the fluctuations in the tail of the degree distribution are very large for the power-law case.

Dorogovtsev *et al.* [126] have also shown that Eq. (13) holds for networks generated using the "preferential attachment" procedure of Barabási and Albert [32] described in Sec. VII.B, and a detailed numerical study of this case has been carried out by Moreira *et al.* [295].

D. Network resilience

Related to degree distributions is the property of resilience of networks to the removal of their vertices, which has been the subject of a good deal of attention in the literature. Most of the networks we have been considering rely for their function on their connectivity, i.e., the existence of paths leading between pairs of vertices. If vertices are removed from a network, the typical length of these paths will increase, and ultimately vertex pairs will become disconnected and communication between them through the network will become impossible. Networks vary in their level of resilience to such vertex removal.

There are also a variety of different ways in which vertices can be removed and different networks show varying degrees of resilience to these also. For example, one could remove vertices at random from a network, or one could target some specific class of vertices, such as those with the highest degrees. Network resilience is of particular importance in epidemiology, where "removal" of vertices in a contact network might correspond for example to vaccination of individuals against a disease. Because vaccination not only prevents the vaccinated individuals from catching the disease but may also destroy paths between other individuals by which the disease might have spread, it can have a wider reaching effect than one might at first think, and careful consideration of the efficacy of different vaccination strategies could lead to substantial advantages for public health.

Recent interest in network resilience has been sparked by the work of Albert *et al.* [15], who studied the effect of vertex deletion in two example networks, a 6000-vertex network representing the topology of the Internet at the level of autonomous systems (see Sec. II.C), and a 326 000-page subset of the World Wide Web. Both of the Internet and the Web have been observed to have degree distributions that are approximately power-law in form [14, 74, 86, 148, 401] (Sec. III.C.1). The authors measured average vertex–vertex distances as a function
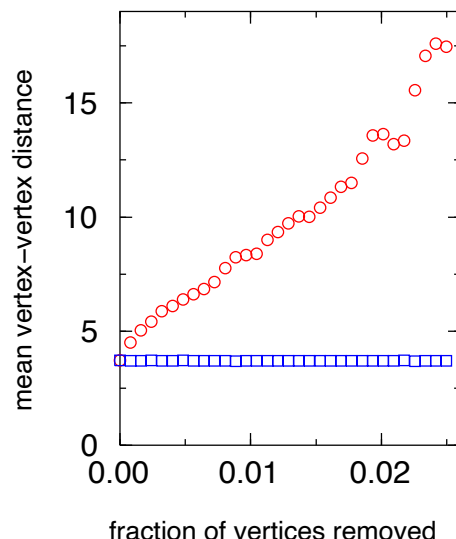

FIG. 7 Mean vertex–vertex distance on a graph representation of the Internet at the autonomous system level, as vertices are removed one by one. If vertices are removed in random order (squares), distance increases only very slightly, but if they are removed in order of their degrees, starting with the highest degree vertices (circles), then distance increases sharply. After Albert *et al.* [15].

of number of vertices removed, both for random removal and for progressive removal of the vertices with the highest degrees.[14] In Fig. 7 we show their results for the Internet. They found for both networks that distance was almost entirely unaffected by random vertex removal, i.e., the networks studied were highly resilient to this type of removal. This is intuitively reasonable, since most of the vertices in these networks have low degree and therefore lie on few paths between others; thus their removal rarely affects communications substantially. On the other hand, when removal is targeted at the highest degree vertices, it is found to have devastating effect. Mean vertex–vertex distance increases very sharply with the fraction of vertices removed, and typically only a few percent of vertices need be removed before essentially all communication through the network is destroyed. Albert *et al.* expressed their results in terms of failure or sabotage of network nodes. The Internet (and the Web) they suggest, is highly resilient against the random failure of vertices in the network, but highly vulnerable to deliberate attack on its highest-degree vertices.

Similar results to those of Albert *et al.* were found independently by Broder *et al.* [74] for a much larger subset of the Web graph. Interestingly, however, Broder *et al.*

---

[14] In removing the vertices with the highest degrees, Albert *et al.* recalculated degrees following the removal of each vertex. Most other authors who have studied this issue have adopted a slightly different strategy of removing vertices in order of their *initial* degree in the network before any removal.

gave an entirely opposite interpretation of their results. They found that in order to destroy connectivity in the Web one has to remove all vertices with degree greater than five, which seems like a drastic attack on the network, given that some vertices have degrees in the thousands. They thus concluded that the network was very *resilient* against targeted attack. In fact however there is not such a conflict between these results as at first appears. Because of the highly skewed degree distribution of the Web, the fraction of vertices with degree greater than five is only a small fraction of all vertices.

Following these studies, many authors have looked into the question of resilience for other networks. In general the picture seems to be consistent with that seen in the Internet and Web. Most networks are robust against random vertex removal but considerably less robust to targeted removal of the highest-degree vertices. Jeong *et al.* [212] have looked at metabolic networks, Dunne *et al.* [132, 133] at food webs, Newman *et al.* [321] at email networks, and a variety of authors at resilience of model networks [15, 81, 93, 94, 200], which we discuss in more detail in later sections of the review. A particularly thorough study of the resilience of both real-world and model networks has been conducted by Holme *et al.* [200], who looked not only at vertex removal but also at removal of edges, and considered some additional strategies for selecting vertices based on so-called "betweenness" (see Secs. III.G and III.J).

E. Mixing patterns

Delving a little deeper into the statistics of network structure, one can ask about which vertices pair up with which others. In most kinds of networks there are at least a few different types of vertices, and the probabilities of connection between vertices often depends on types. For example, in a food web representing which species eat which in an ecosystem (Sec. II.D) one sees vertices representing plants, herbivores, and carnivores. Many edges link the plants and herbivores, and many more the herbivores and carnivores. But there are few edges linking herbivores to other herbivores, or carnivores to plants. For the Internet, Maslov *et al.* [275] have proposed that the structure of the network reflects the existence of three broad categories of nodes: high-level connectivity providers who run the Internet backbone and trunk lines, consumers who are end users of Internet service, and ISPs who join the two. Again there are many links between end users and ISPs, and many between ISPs and backbone operators, but few between ISPs and other ISPs, or between backbone operators and end users.

In social networks this kind of selective linking is called *assortative mixing* or *homophily* and has been widely studied, as it has also in epidemiology. (The term "assortative matching" is also seen in the ecology literature, particularly in reference to mate choice among animals.)

		women			
		black	hispanic	white	other
men	black	506	32	69	26
	hispanic	23	308	114	38
	white	26	46	599	68
	other	10	14	47	32

TABLE III Couples in the study of Catania *et al.* [85] tabulated by race of either partner. After Morris [302].

A classic example of assortative mixing in social networks is mixing by race. Table III for example reproduces results from a study of 1 958 couples in the city of San Francisco, California. Among other things, the study recorded the race (self-identified) of study participants in each couple. As the table shows, participants appear to draw their partners preferentially from those of their own race, and this is believed to be a common phenomenon in many social networks: we tend to associate preferentially with people who are similar to ourselves in some way.

Assortative mixing can be quantified by an "assortativity coefficient," which can be defined in a couple of different ways. Let E_{ij} be the number of edges in a network that connect vertices of types i and j, with $i, j = 1 \ldots N$, and let $\mathbf{E}$ be the matrix with elements E_{ij}, as depicted in Table III. We define a normalized mixing matrix by

$$\mathbf{e} = \frac{\mathbf{E}}{\| \mathbf{E} \|}, \tag{14}$$

where $\| \mathbf{x} \|$ means the sum of all the elements of the matrix $\mathbf{x}$. The elements e_{ij} measure the *fraction* of edges that fall between vertices of types i and j. One can also ask about the conditional probability $P(j|i)$ that my network neighbor is of type j given that I am of type i, which is given by $P(j|i) = e_{ij}/\sum_j e_{ij}$. These quantities satisfy the normalization conditions

$$\sum_{ij} e_{ij} = 1, \qquad \sum_j P(j|i) = 1. \tag{15}$$

Gupta *et al.* [186] have suggested that assortative mixing be quantified by the coefficient

$$Q = \frac{\sum_i P(i|i) - 1}{N - 1}. \tag{16}$$

This quantity has the desirable properties that it is 1 for a perfectly assortative network (every edge falls between vertices of the same type), and 0 for randomly mixed networks, and it has been quite widely used in the literature. But it suffers from two shortcomings [318]: (1) for an asymmetric matrix like the one in Table III, Q has two different values, depending on whether we put the men or the women along the horizontal axis, and it is unclear which of these two values is the "correct" one for the network; (2) the measure weights each vertex type equally, regardless of how many vertices there are of each type,

which can give rise to misleading figures for Q in cases where community size is heterogeneous, as it often is.

An alternative assortativity coefficient that remedies these problems is defined by [318]

$$r = \frac{\text{Tr}\,\mathbf{e} - \| \mathbf{e}^2 \|}{1 - \| \mathbf{e}^2 \|}. \tag{17}$$

This quantity is also 0 in a randomly mixed network and 1 in a perfectly assortative one. But its value is not altered by transposition of the matrix and it weights vertices equally rather than communities, so that small communities make an appropriately small contribution to r. For the data of Table III we find $r = 0.621$.

Another type of assortative mixing is mixing by scalar characteristics such as age or income. Again it is usually found that people prefer to associate with others of similar age and income to themselves, although of course age and income, like race, may be proxies for other driving forces, such as cultural differences. Garfinkel et al. [170] and Newman [318], for example, have analyzed data for unmarried and married couples respectively to show that there is strong correlation between the ages of partners. Mixing by scalar characteristics can be quantified by calculating a correlation coefficient for the characteristic in question.

In theory assortative mixing according to vector characteristics should also be possible. For example, geographic location probably affects individuals' propensity to become acquainted. Location could be viewed as a two-vector, with the probability of connection between pairs of individuals being assortative on the values of these vectors.

F. Degree correlations

A special case of assortative mixing according to a scalar vertex property is mixing according to vertex degree, also commonly referred to simply as degree correlation. Do the high-degree vertices in a network associate preferentially with other high-degree vertices? Or do they prefer to attach to low-degree ones? Both situations are seen in some networks, as it turns out. The case of assortative mixing by degree is of particular interest because, since degree is itself a property of the graph topology, degree correlations can give rise to some interesting network structure effects.

Several different ways of quantifying degree correlations have been proposed. Maslov et al. [274, 275] have simply plotted the two-dimensional histogram of the degrees of vertices at either ends of an edge. They have shown results for protein interaction networks and the Internet. A more compact representation of the situation is that proposed by Pastor-Satorras et al. [331, 401], who in studies of the Internet calculated the mean degree of the network neighbors of a vertex as a function of the degree k of that vertex. This gives a one-parameter curve which increases with k if the network is assortatively mixed. For the Internet in fact it is found to decrease with k, a situation we call disassortativity. Newman [314, 318] reduced the measurement still further to a single number by calculating the Pearson correlation coefficient of the degrees at either ends of an edge. This gives a single number that should be positive for assortatively mixed networks and negative for disassortative ones. In Table II we show results for a number of different networks. An interesting observation is that essentially all social networks measured appear to be assortative, but other types of networks (information networks, technological networks, biological networks) appear to be disassortative. It is not clear what the explanation for this result is, or even if there is any one single explanation. (Probably there is not.)

G. Community structure

It is widely assumed [363, 409] that most social networks show "community structure," i.e., groups of vertices that have a high density of edges within them, with a lower density of edges between groups. It is a matter of common experience that people do divide into groups along lines of interest, occupation, age, and so forth, and the phenomenon of assortativity discussed in Sec. III.F certainly suggests that this might be the case. (It is possible for a network to have assortative mixing but no community structure. This can occur, for example, when there is assortative mixing by age or other scalar quantities. Networks with this type of structure are sometimes said to be "stratified.")

In Fig. 8 we show a visualization of the friendship network of children in a US school taken from a study by Moody [291].[15] The figure was created using a "spring embedding" algorithm, in which linear springs are placed between vertices and the system is relaxed using a first-order energy minimization. We have no special reason to suppose that this very simple algorithm would reveal anything particularly useful about the network, but the network appears to have strong enough community structure that in fact the communities appear clearly in the figure. Moreover, when Moody colors the vertices according to the race of the individuals they represent, as shown in the figure, it becomes immediately clear that one of the principal divisions in the network is by individuals' race, and this is presumably what is driving the formation of communities in this case. (The other principal division visible in the figure is between middle school and high school, which are age divisions in the American education system.)

---

[15] This image does not appear in the paper cited, but it and a number of other images from the same study can be found on the Web at http://www.sociology.ohio-state.edu/jwm/.
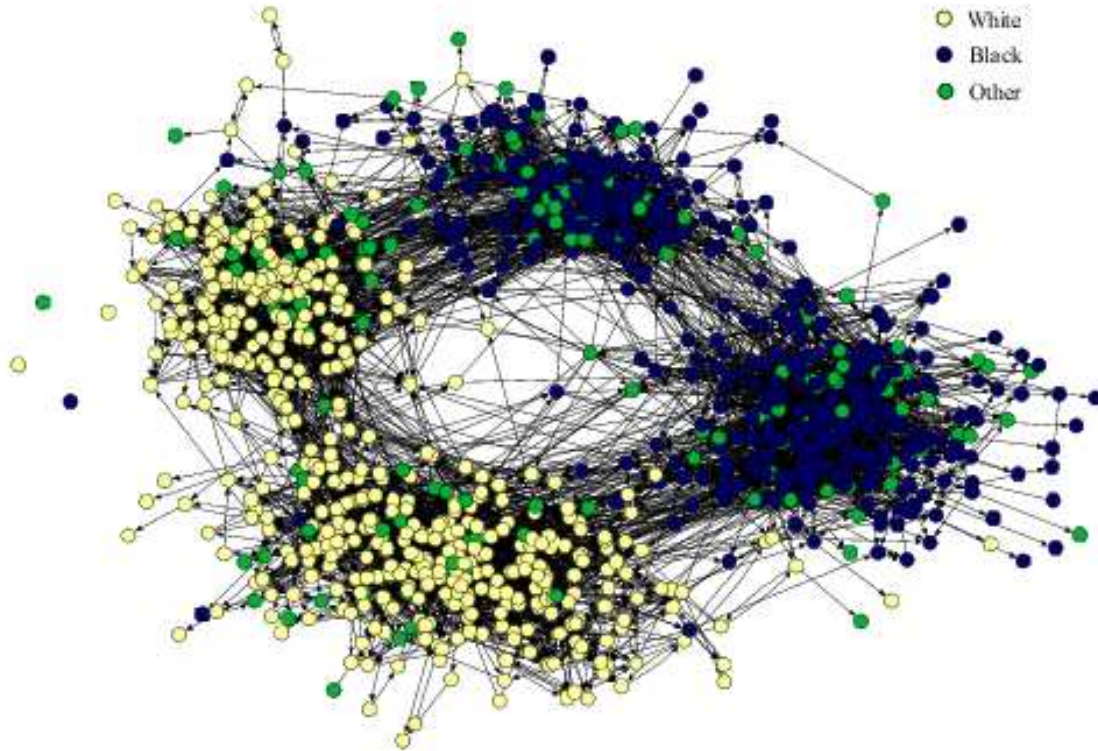
FIG. 8 Friendship network of children in a US school. Friendships are determined by asking the participants, and hence are directed, since A may say that B is their friend but not *vice versa*. Vertices are color coded according to race, as marked, and the split from left to right in the figure is clearly primarily along lines of race. The split from top to bottom is between middle school and high school, i.e., between younger and older children. Picture courtesy of James Moody.

It would be of some interest, and indeed practical importance, were we to find that other types of networks, such as those those listed in Table II, show similar group structure also. One might well imagine for example that citation networks would divide into groups representing particular areas of research interest, and a good deal of energy has been invested in studies of this phenomenon [101, 138]. Similarly communities in the World Wide Web might reflect the subject matter of pages, communities in metabolic, neural, or software networks might reflect functional units, communities in food webs might reflect subsystems within ecosystems, and so on.

The traditional method for extracting community structure from a network is *cluster analysis* [147], sometimes also called *hierarchical clustering*.[16] In this method, one assigns a "connection strength" to vertex pairs in the network of interest. In general each of the $\frac{1}{2}n(n-1)$ possible pairs in a network of n vertices is assigned such a strength, not just those that are connected by an edge, although there are versions of the

method where not all pairs are assigned a strength; in that case one can assume the remaining pairs to have a connection strength of zero. Then, starting with n vertices with no edges between any of them, one adds edges in order of decreasing vertex–vertex connection strength. One can pause at any point in this process and examine the component structure formed by the edges added so far; these components are taken to be the communities (or "clusters") at that stage in the process. When all edges have been added, all vertices are connected to all others, and there is only one community. The entire process can be represented by a tree or *dendrogram* of union operations between vertex sets in which the communities at any level correspond to a horizontal cut through the tree—see Fig. 9.[17]

Clustering is possible according to many different definitions of the connection strength. Reasonable choices include various weighted vertex–vertex distance measures, the sizes of minimum cut-sets (i.e., maximum flow) [7],

---

[16] Not to be confused with the entirely different use of the word clustering introduced in Sec. III.B.

[17] For some reason such trees are conventionally depicted with their "root" at the top and their "leaves" at the bottom, which is not the natural order of things for most trees.
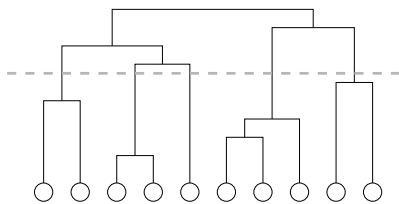
FIG. 9 An example of a dendrogram showing the hierarchical clustering of ten vertices. A horizontal cut through the dendrogram, such as that denoted by the dotted line, splits the vertices into a set of communities, five in this case.

and weighted path counts between vertices. Recently a number of authors have had success with methods based on "edge betweenness," which is the count of how many geodesic paths between vertices run along each edge in the network [171, 185, 197, 422]. Results appear to show that, for social and biological networks at least, community structure is a common network property, although some food webs are found not to break up into communities in any simple way. (Food webs may be different from other networks in that they appear to be dense: mean vertex degree increases roughly linearly with network size, rather than remaining constant as it does in most networks [132, 273]. The same may be true of metabolic networks also [P. Holme, personal communication].)

Network clustering should not be confused with the technique of data clustering, which is a way of detecting groupings of data-points in high-dimensional data spaces [208]. The two problems do have some common features however, and algorithms for one can be adapted for the other, and *vice versa*. For example, high-dimensional data can be converted into a network by placing edges between closely spaced data points, and then network clustering algorithms can be applied to the result. On balance, however, one normally finds that algorithms specially devised for data clustering work better than such borrowed methods, and the same is true in reverse.

In the social networks literature, network clustering has been discussed to a great extent in the context of so-called *block models*, [71, 419] which are essentially just divisions of networks into communities or blocks according to one criterion or another. Sociologists have concentrated particularly on *structural equivalence*. Two vertices in a network are said to be structurally equivalent if they have all of the same neighbors. Exact structural equivalence is rare, but approximate equivalence can be used as the basis for a hierarchical clustering method such as that described above.

Another slightly different question about community structure, but related to the one discussed here, has been studied by Flake *et al.* [158]: if one is given an example vertex drawn from a known network, can one identify the community to which it belongs? Algorithmic methods for answering this question would clearly be of some practical value for searching networks such as the World Wide Web and citation networks. Flake *et al.* give what appears to be a very successful algorithm, at least in the context of the Web, based on a maximum flow method.

H. Network navigation

Stanley Milgram's famous small-world experiment (Sec. II.A), in which letters were passed from person to person in an attempt to get them to a desired target individual, showed that there exist short paths through social networks between apparently distant individuals. However, there is another conclusion that can be drawn from this experiment which Milgram apparently failed to notice; it was pointed out in 2000 by Kleinberg [238, 239]. Milgram's results demonstrate that there exist short paths in the network, but they also demonstrate that ordinary people are good at finding them. This is, upon reflection, perhaps an even more surprising result than the existence of the paths in the first place. The participants in Milgram's study had no special knowledge of the network connecting them to the target person. Most people know only who their friends are and perhaps a few of their friends' friends. Nonetheless it proved possible to get a message to a distant target in only a small number of steps. This indicates that there is something quite special about the structure of the network. On a random graph for instance, as Kleinberg pointed out, short paths between vertices exist but no one would be able to find them given only the kind of information that people have in realistic situations. If it were possible to construct artificial networks that were easy to navigate in the same way that social networks appear to be, it has been suggested they could be used to build efficient database structures or better peer-to-peer computer networks [5, 6, 415] (see Sec. VIII.C.3).

I. Other network properties

In addition to the heavily studied network properties of the preceding sections, a number of others have received some attention. In some networks the size of the largest component is an important quantity. For example, in a communication network like the Internet the size of the largest component represents the largest fraction of the network within which communication is possible and hence is a measure of the effectiveness of the network at doing its job [74, 81, 93, 94, 125, 323]. The size of the largest component is often equated with the graph theoretical concept of the "giant component" (see Sec. IV.A), although technically the two are only the same in the limit of large graph size. The size of the second-largest component in a network is also measured sometimes. In networks well above the density at which a giant component first forms, the largest component is expected to be much larger than the second largest (Sec. IV.A).

if these are the only important quantities to measure (almost certainly they are not) or even if they are the most important. We have as yet no theoretical framework to tell us if we are even looking in the right place. Perhaps there are other measures, so far un-thought-of, that are more important than those we have at present. A true understanding of which properties of networks are the important ones to focus on will almost certainly require us to state first what questions we are interested in answering about a particular network. And knowing how to tie the answers to these questions to structural properties of the network is therefore also an important goal.

Second, there is much to be done in developing more sophisticated models of networks, both to help us understand network topology and to act as a substrate for the study of processes taking place on networks. While some network properties, such as degree distributions,

have been thoroughly modeled and their causes and effects well understood, others such as correlations, transitivity, and community structure have not. It seems certain that these properties will affect the behavior of networked systems substantially, so our current lack of suitable techniques to handle them leaves a large gap in our understanding.

Which leads us to our third and perhaps most important direction for future study, the behavior of processes taking place on networks. The work described in Sec. VIII represents only a few first attempts at answering questions about such processes, and yet this, in a sense, is our ultimate goal in this field: to understand the behavior and function of the networked systems we see around us. If we can gain such understanding, it will give us new insight into a vast array of complex and previously poorly understood phenomena.

References

[1] Abramson, G. and Kuperman, M., Social games in a social network, *Phys. Rev. E* **63**, 030901 (2001).

[2] Adamic, L. A., The small world web, in *Lecture Notes in Computer Science*, vol. 1696, pp. 443–454, Springer, New York (1999).

[3] Adamic, L. A. and Adar, E., Friends and neighbors on the Web, *Social Networks* (in press).

[4] Adamic, L. A. and Huberman, B. A., Power-law distribution of the world wide web, *Science* **287**, 2115 (2000).

[5] Adamic, L. A., Lukose, R. M., and Huberman, B. A., Local search in unstructured networks, in S. Bornholdt and H. G. Schuster (eds.), *Handbook of Graphs and Networks*, Wiley-VCH, Berlin (2003).

[6] Adamic, L. A., Lukose, R. M., Puniyani, A. R., and Huberman, B. A., Search in power-law networks, *Phys. Rev. E* **64**, 046135 (2001).

[7] Ahuja, R. K., Magnanti, T. L., and Orlin, J. B., *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall, Upper Saddle River, New Jersey (1993).

[8] Aiello, W., Chung, F., and Lu, L., A random graph model for massive graphs, in *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, pp. 171–180, Association of Computing Machinery, New York (2000).

[9] Aiello, W., Chung, F., and Lu, L., Random evolution of massive graphs, in J. Abello, P. M. Pardalos, and M. G. C. Resende (eds.), *Handbook of Massive Data Sets*, pp. 97–122, Kluwer, Dordrecht (2002).

[10] Alberich, R., Miro-Julia, J., and Rossello, F., Marvel Universe looks almost like a real social network, Preprint cond-mat/0202174 (2002).

[11] Albert, R. and Barabási, A.-L., Dynamics of complex systems: Scaling laws for the period of Boolean networks, *Phys. Rev. Lett.* **84**, 5660–5663 (2000).

[12] Albert, R. and Barabási, A.-L., Topology of evolving

networks: Local events and universality, *Phys. Rev. Lett.* **85**, 5234–5237 (2000).

[13] Albert, R. and Barabási, A.-L., Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74**, 47–97 (2002).

[14] Albert, R., Jeong, H., and Barabási, A.-L., Diameter of the world-wide web, *Nature* **401**, 130–131 (1999).

[15] Albert, R., Jeong, H., and Barabási, A.-L., Attack and error tolerance of complex networks, *Nature* **406**, 378–382 (2000).

[16] Aldana, M., Dynamics of Boolean networks with scale-free topology, Preprint cond-mat/0209571 (2002).

[17] Aldous, D. and Pittel, B., On a random graph with immigrating vertices: Emergence of the giant component, *Random Structures and Algorithms* **17**, 79–102 (2000).

[18] Aleksiejuk, A., Hołyst, J. A., and Stauffer, D., Ferromagnetic phase transition in Barabási–Albert networks, *Physica A* **310**, 260–266 (2002).

[19] Almaas, E., Kulkarni, R. V., and Stroud, D., Characterizing the structure of small-world networks, *Phys. Rev. Lett.* **88**, 098101 (2002).

[20] Amaral, L. A. N., Scala, A., Barthélémy, M., and Stanley, H. E., Classes of small-world networks, *Proc. Natl. Acad. Sci. USA* **97**, 11149–11152 (2000).

[21] Ancel Meyers, L., Newman, M. E. J., Martin, M., and Schrag, S., Applying network theory to epidemics: Control measures for outbreaks of Mycoplasma pneumoniae, *Emerging Infectious Diseases* **9**, 204–210 (2001).

[22] Anderson, C., Wasserman, S., and Crouch, B., A p* primer: Logit models for social networks, *Social Networks* **21**, 37–66 (1999).

[23] Anderson, R. M. and May, R. M., *Infectious Diseases of Humans*, Oxford University Press, Oxford (1991).

[24] Andersson, H., Epidemic models and social networks, *Math. Scientist* **24**, 128–147 (1999).

[25] Arenas, A., Cabrales, A., Díaz-Guilera, A., Guimerà, R., and Vega-Redondo, F., Search and congestion in complex networks, in R. Pastor-Satorras and J. Rubi

(eds.), *Proceedings of the XVIII Sitges Conference on Statistical Mechanics*, Lecture Notes in Physics, Springer, Berlin (2003).

[26] Bailey, N. T. J., *The Mathematical Theory of Infectious Diseases and Its Applications*, Hafner Press, New York (1975).

[27] Baird, D. and Ulanowicz, R. E., The seasonal dynamics of the Chesapeake Bay ecosystem, *Ecological Monographs* **59**, 329–364 (1989).

[28] Ball, F., Mollison, D., and Scalia-Tomba, G., Epidemics with two levels of mixing, *Annals of Applied Probability* **7**, 46–89 (1997).

[29] Banavar, J. R., Maritan, A., and Rinaldo, A., Size and form in efficient transportation networks, *Nature* **399**, 130–132 (1999).

[30] Banks, D. L. and Carley, K. M., Models for network evolution, *Journal of Mathematical Sociology* **21**, 173–196 (1996).

[31] Barabási, A.-L., *Linked: The New Science of Networks*, Perseus, Cambridge, MA (2002).

[32] Barabási, A.-L. and Albert, R., Emergence of scaling in random networks, *Science* **286**, 509–512 (1999).

[33] Barabási, A.-L., Albert, R., and Jeong, H., Mean-field theory for scale-free random networks, *Physica A* **272**, 173–187 (1999).

[34] Barabási, A.-L., Albert, R., and Jeong, H., Scale-free characteristics of random networks: The topology of the World Wide Web, *Physica A* **281**, 69–77 (2000).

[35] Barabási, A.-L., Albert, R., Jeong, H., and Bianconi, G., Power-law distribution of the World Wide Web, *Science* **287**, 2115a (2000).

[36] Barabási, A.-L., Jeong, H., Ravasz, E., Néda, Z., Schuberts, A., and Vicsek, T., Evolution of the social network of scientific collaborations, *Physica A* **311**, 590–614 (2002).

[37] Barahona, M. and Pecora, L. M., Synchronization in small-world systems, *Phys. Rev. Lett.* **89**, 054101 (2002).

[38] Barbour, A. D. and Reinert, G., Small worlds, *Random Structures and Algorithms* **19**, 54–74 (2001).

[39] Barrat, A., Comment on 'Small-world networks: Evidence for crossover picture', Preprint cond-mat/9903323 (1999).

[40] Barrat, A. and Weigt, M., On the properties of small-world networks, *Eur. Phys. J. B* **13**, 547–560 (2000).

[41] Barthélémy, M. and Amaral, L. A. N., Erratum: Small-world networks: Evidence for a crossover picture, *Phys. Rev. Lett.* **82**, 5180 (1999).

[42] Barthélémy, M. and Amaral, L. A. N., Small-world networks: Evidence for a crossover picture, *Phys. Rev. Lett.* **82**, 3180–3183 (1999).

[43] Batagelj, V. and Mrvar, A., Some analyses of Erdős collaboration graph, *Social Networks* **22**, 173–186 (2000).

[44] Bauer, M. and Bernard, D., A simple asymmetric evolving random network, Preprint cond-mat/0203232 (2002).

[45] Bearman, P. S., Moody, J., and Stovel, K., Chains of affection: The structure of adolescent romantic and sexual networks, Preprint, Department of Sociology, Columbia University (2002).

[46] Bekessy, A., Bekessy, P., and Komlos, J., Asymptotic enumeration of regular matrices, *Stud. Sci. Math. Hungar.* **7**, 343–353 (1972).

[47] Bender, E. A. and Canfield, E. R., The asymptotic number of labeled graphs with given degree sequences, *Journal of Combinatorial Theory A* **24**, 296–307 (1978).

[48] Berg, J. and Lässig, M., Correlated random networks, *Phys. Rev. Lett.* **89**, 228701 (2002).

[49] Berg, J., Lässig, M., and Wagner, A., Evolution dynamics of protein networks, Preprint cond-mat/0207711 (2002).

[50] Bernard, H. R., Killworth, P. D., Evans, M. J., McCarty, C., and Shelley, G. A., Studying social relations cross-culturally, *Ethnology* **2**, 155–179 (1988).

[51] Bianconi, G., Mean field solution of the Ising model on a Barabási–Albert network, Preprint cond-mat/0204455 (2002).

[52] Bianconi, G. and Barabási, A.-L., Bose–Einstein condensation in complex networks, *Phys. Rev. Lett.* **86**, 5632–5635 (2001).

[53] Bianconi, G. and Barabási, A.-L., Competition and multiscaling in evolving networks, *Europhys. Lett.* **54**, 436–442 (2001).

[54] Bianconi, G. and Capocci, A., Number of loops of size h in growing scale-free networks, *Phys. Rev. Lett.* **90**, 078701 (2003).

[55] Bilke, S. and Peterson, C., Topological properties of citation and metabolic networks, *Phys. Rev. E* **64**, 036106 (2001).

[56] Blanchard, P., Chang, C.-H., and Krüger, T., Epidemic thresholds on scale-free graphs: The interplay between exponent and preferential choice, Preprint cond-mat/0207319 (2002).

[57] Boguñá, M. and Pastor-Satorras, R., Epidemic spreading in correlated complex networks, *Phys. Rev. E* **66**, 047104 (2002).

[58] Boguñá, M., Pastor-Satorras, R., and Vespignani, A., Absence of epidemic threshold in scale-free networks with connectivity correlations, Preprint cond-mat/0208163 (2002).

[59] Boguñá, M., Pastor-Satorras, R., and Vespignani, A., Epidemic spreading in complex networks with degree correlations, in R. Pastor-Satorras and J. Rubi (eds.), *Proceedings of the XVIII Sitges Conference on Statistical Mechanics*, Lecture Notes in Physics, Springer, Berlin (2003).

[60] Bollobás, B., A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European Journal on Combinatorics* **1**, 311–316 (1980).

[61] Bollobás, B., The diameter of random graphs, *Trans. Amer. Math. Soc.* **267**, 41–52 (1981).

[62] Bollobás, B., *Modern Graph Theory*, Springer, New York (1998).

[63] Bollobás, B., *Random Graphs*, Academic Press, New York, 2nd ed. (2001).

[64] Bollobás, B. and Riordan, O., The diameter of a scale-free random graph, Preprint, Department of Mathematical Sciences, University of Memphis (2002).

[65] Bollobás, B., Riordan, O., Spencer, J., and Tusnády, G., The degree sequence of a scale-free random graph process, *Random Structures and Algorithms* **18**, 279–290 (2001).

[66] Bonacich, P. F., A technique for analyzing overlapping memberships, in H. Costner (ed.), *Sociological Methodology*, Jossey-Bass, San Francisco (1972).

[67] Bonacich, P. F., Power and centrality: A family of measures, *Am. J. Sociol.* **92**, 1170–1182 (1987).

[68] Bordens, M. and Gómez, I., Collaboration networks

in science, in H. B. Atkins and B. Cronin (eds.), *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield*, Information Today, Medford, NJ (2000).

[69] Bornholdt, S. and Ebel, H., World Wide Web scaling exponent from Simon's 1955 model, *Phys. Rev. E* **64**, 035104 (2001).

[70] Bornholdt, S. and Schuster, H. G. (eds.), *Handbook of Graphs and Networks*, Wiley-VCH, Berlin (2003).

[71] Breiger, R. L., Boorman, S. A., and Arabie, P., An algorithm for clustering relations data with applications to social network analysis and comparison with multidimensional scaling, *Journal of Mathematical Psychology* **12**, 328–383 (1975).

[72] Brin, S. and Page, L., The anatomy of a large-scale hypertextual Web search engine, *Computer Networks* **30**, 107–117 (1998).

[73] Broadbent, S. R. and Hammersley, J. M., Percolation processes: I. Crystals and mazes, *Proc. Cambridge Philos. Soc.* **53**, 629–641 (1957).

[74] Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., and Wiener, J., Graph structure in the web, *Computer Networks* **33**, 309–320 (2000).

[75] Broida, A. and Claffy, K. C., Internet topology: Connectivity of IP graphs, in S. Fahmy and K. Park (eds.), *Scalability and Traffic Control in IP Networks*, no. 4526 in Proc. SPIE, pp. 172–187, International Society for Optical Engineering, Bellingham, WA (2001).

[76] Buchanan, M., *Nexus: Small Worlds and the Groundbreaking Science of Networks*, Norton, New York (2002).

[77] Burda, Z., Correia, J. D., and Krzywicki, A., Statistical ensemble of scale-free random graphs, *Phys. Rev. E* **64**, 046118 (2001).

[78] Caldarelli, G., Capocci, A., De Los Rios, P., and Muñoz, M. A., Scale-free networks from varying vertex intrinsic fitness, *Phys. Rev. Lett.* **89**, 258702 (2002).

[79] Caldarelli, G., Pastor-Satorras, R., and Vespignani, A., Cycles structure and local ordering in complex networks, Preprint cond-mat/0212026 (2002).

[80] Callaway, D. S., Hopcroft, J. E., Kleinberg, J. M., Newman, M. E. J., and Strogatz, S. H., Are randomly grown graphs really random?, *Phys. Rev. E* **64**, 041902 (2001).

[81] Callaway, D. S., Newman, M. E. J., Strogatz, S. H., and Watts, D. J., Network robustness and fragility: Percolation on random graphs, *Phys. Rev. Lett.* **85**, 5468–5471 (2000).

[82] Camacho, J., Guimerà, R., and Amaral, L. A. N., Robust patterns in food web structure, *Phys. Rev. Lett.* **88**, 228102 (2002).

[83] Capocci, A., Caldarelli, G., and De Los Rios, P., Quantitative description and modeling of real networks, Preprint cond-mat/0206336 (2002).

[84] Castellano, C., Vilone, D., and Vespignani, A., Incomplete ordering of the voter model on small-world networks, Preprint cond-mat/0210465 (2002).

[85] Catania, J. A., Coates, T. J., Kegels, S., and Fullilove, M. T., The population-based AMEN (AIDS in Multi-Ethnic Neighborhoods) study, *Am. J. Public Health* **82**, 284–287 (1992).

[86] Chen, Q., Chang, H., Govindan, R., Jamin, S., Shenker, S. J., and Willinger, W., The origin of power laws in Internet topologies revisited, in *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies*, IEEE Computer Soci-

ety (2002).

[87] Chowell, G., Hyman, J. M., and Eubank, S., Analysis of a real world network: The City of Portland, Technical Report BU-1604-M, Department of Biological Statistics and Computational Biology, Cornell University (2002).

[88] Chung, F. and Lu, L., The average distances in random graphs with given expected degrees, *Proc. Natl. Acad. Sci. USA* **99**, 15879–15882 (2002).

[89] Chung, F. and Lu, L., Connected components in random graphs with given degree sequences, *Annals of Combinatorics* **6**, 125–145 (2002).

[90] Chung, F., Lu, L., Dewey, T. G., and Galas, D. J., Duplication models for biological networks, *Journal of Computational Biology* (in press).

[91] Cohen, J. E., Briand, F., and Newman, C. M., *Community food webs: Data and theory*, Springer, New York (1990).

[92] Cohen, R., ben-Avraham, D., and Havlin, S., Efficient immunization of populations and computers, Preprint cond-mat/0207387 (2002).

[93] Cohen, R., Erez, K., ben-Avraham, D., and Havlin, S., Resilience of the Internet to random breakdowns, *Phys. Rev. Lett.* **85**, 4626–4628 (2000).

[94] Cohen, R., Erez, K., ben-Avraham, D., and Havlin, S., Breakdown of the Internet under intentional attack, *Phys. Rev. Lett.* **86**, 3682–3685 (2001).

[95] Cohen, R. and Havlin, S., Scale-free networks are ultrasmall, *Phys. Rev. Lett.* **90**, 058701 (2003).

[96] Connor, R. C., Heithaus, M. R., and Barre, L. M., Superalliance of bottlenose dolphins, *Nature* **397**, 571–572 (1999).

[97] Coppersmith, S. N., Kadanoff, L. P., and Zhang, Z., Reversible Boolean networks: I. Distribution of cycle lengths, *Physica D* **149**, 11–29 (2000).

[98] Coppersmith, S. N., Kadanoff, L. P., and Zhang, Z., Reversible Boolean networks: II. Phase transition, oscillation, and local structures, *Physica D* **157**, 54–74 (2001).

[99] Corman, S. R., Kuhn, T., Mcphee, R. D., and Dooley, K. J., Studying complex discursive systems: Centering resonance analysis of organizational communication, *Human Communication Research* **28**, 157–206 (2002).

[100] Coulumb, S. and Bauer, M., Asymmetric evolving random networks, Preprint cond-mat/0212371 (2002).

[101] Crane, D., *Invisible colleges: Diffusion of knowledge in scientific communities*, University of Chicago Press, Chicago (1972).

[102] Davidsen, J., Ebel, H., , and Bornholdt, S., Emergence of a small world from local interactions: Modeling acquaintance networks, *Phys. Rev. Lett.* **88**, 128701 (2002).

[103] Davis, A., Gardner, B. B., and Gardner, M. R., *Deep South*, University of Chicago Press, Chicago (1941).

[104] Davis, G. F. and Greve, H. R., Corporate elite networks and governance changes in the 1980s, *Am. J. Sociol.* **103**, 1–37 (1997).

[105] Davis, G. F., Yoo, M., and Baker, W. E., The small world of the corporate elite, Preprint, University of Michigan Business School (2001).

[106] de Arcangelis, L. and Herrmann, H. J., Self-organized criticality on small world networks, *Physica A* **308**, 545–549 (2002).

[107] de Castro, R. and Grossman, J. W., Famous trails to Paul Erdős, *Mathematical Intelligencer* **21**, 51–63

(1999).

[108] de Groot, M. H., Reaching a consensus, *J. Amer. Statist. Assoc.* **69**, 118–121 (1974).

[109] de Menezes, M. A., Moukarzel, C., and Penna, T. J. P., First-order transition in small-world networks, *Europhys. Lett.* **50**, 574–579 (2000).

[110] Dodds, P. S., Muhamad, R., and Watts, D. J., An experiment study of social search and the small world problem, Preprint, Department of Sociology, Columbia University (2002).

[111] Dodds, P. S. and Rothman, D. H., Geometry of river networks, *Phys. Rev. E* **63**, 016115, 016116, & 016117 (2001).

[112] Dorogovtsev, S. N., Goltsev, A. V., and Mendes, J. F. F., Ising model on networks with an arbitrary distribution of connections, *Phys. Rev. E* **66**, 016104 (2002).

[113] Dorogovtsev, S. N., Goltsev, A. V., and Mendes, J. F. F., Pseudofractal scale-free web, *Phys. Rev. E* **65**, 066122 (2002).

[114] Dorogovtsev, S. N. and Mendes, J. F. F., Evolution of networks with aging of sites, *Phys. Rev. E* **62**, 1842–1845 (2000).

[115] Dorogovtsev, S. N. and Mendes, J. F. F., Exactly solvable small-world network, *Europhys. Lett.* **50**, 1–7 (2000).

[116] Dorogovtsev, S. N. and Mendes, J. F. F., Scaling behaviour of developing and decaying networks, *Europhys. Lett.* **52**, 33–39 (2000).

[117] Dorogovtsev, S. N. and Mendes, J. F. F., Comment on "Breakdown of the Internet under intentional attack", *Phys. Rev. Lett.* **87**, 219801 (2001).

[118] Dorogovtsev, S. N. and Mendes, J. F. F., Effect of the accelerating growth of communications networks on their structure, *Phys. Rev. E* **63**, 025101 (2001).

[119] Dorogovtsev, S. N. and Mendes, J. F. F., Language as an evolving word web, *Proc. R. Soc. London B* **268**, 2603–2606 (2001).

[120] Dorogovtsev, S. N. and Mendes, J. F. F., Evolution of networks, *Advances in Physics* **51**, 1079–1187 (2002).

[121] Dorogovtsev, S. N. and Mendes, J. F. F., Accelerated growth of networks, in S. Bornholdt and H. G. Schuster (eds.), *Handbook of Graphs and Networks*, pp. 318–341, Wiley-VCH, Berlin (2003).

[122] Dorogovtsev, S. N. and Mendes, J. F. F., *Evolution of Networks: From Biological Nets to the Internet and WWW*, Oxford University Press, Oxford (2003).

[123] Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N., Structure of growing networks with preferential linking, *Phys. Rev. Lett.* **85**, 4633–4636 (2000).

[124] Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N., Anomalous percolation properties of growing networks, *Phys. Rev. E* **64**, 066110 (2001).

[125] Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N., Giant strongly connected component of directed networks, *Phys. Rev. E* **64**, 025101 (2001).

[126] Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N., Size-dependent degree distribution of a scale-free growing network, *Phys. Rev. E* **63**, 062101 (2001).

[127] Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N., Metric structure of random networks, Preprint cond-mat/0210085 (2002).

[128] Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N., Principles of statistical mechanics of random networks, Preprint cond-mat/0204111 (2002).

[129] Dorogovtsev, S. N. and Samukhin, A. N., Mesoscopics and fluctuations in networks, Preprint cond-mat/0211518 (2002).

[130] Doye, J. P. K., Network topology of a potential energy landscape: A static scale-free network, *Phys. Rev. Lett.* **88**, 238701 (2002).

[131] Du, D., Gu, J., and Pardalos, P. M. (eds.), *Satisfiability Problem: Theory and Applications*, no. 35 in DIMACS Series in Discrete Mathematics and Theoretical Computer Science, American Mathematical Society, Providence, RI (1997).

[132] Dunne, J. A., Williams, R. J., and Martinez, N. D., Food-web structure and network theory: The role of connectance and size, *Proc. Natl. Acad. Sci. USA* **99**, 12917–12922 (2002).

[133] Dunne, J. A., Williams, R. J., and Martinez, N. D., Network structure and biodiversity loss in food webs: Robustness increases with connectance, *Ecology Letters* **5**, 558–567 (2002).

[134] Durrett, R. T., Rigorous results for the Callaway–Hopcroft–Kleinberg–Newman–Strogatz model, Preprint, Department of Mathematics, Cornell University (2003).

[135] Ebel, H. and Bornholdt, S., Co-evolutionary games on networks, *Phys. Rev. E* **66**, 056118 (2002).

[136] Ebel, H., Mielsch, L.-I., and Bornholdt, S., Scale-free topology of e-mail networks, *Phys. Rev. E* **66**, 035103 (2002).

[137] Eckmann, J.-P. and Moses, E., Curvature of co-links uncovers hidden thematic layers in the world wide web, *Proc. Natl. Acad. Sci. USA* **99**, 5825–5829 (2002).

[138] Egghe, L. and Rousseau, R., *Introduction to Informetrics*, Elsevier, Amsterdam (1990).

[139] Eguiluz, V. M. and Klemm, K., Epidemic threshold in structured scale-free networks, *Phys. Rev. Lett.* **89**, 108701 (2002).

[140] Eigen, M. and Schuster, P., *The Hypercycle: A Principle of Natural Self-Organization*, Springer, New York (1979).

[141] Erdős, P. and Rényi, A., On random graphs, *Publicationes Mathematicae* **6**, 290–297 (1959).

[142] Erdős, P. and Rényi, A., On the evolution of random graphs, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* **5**, 17–61 (1960).

[143] Erdős, P. and Rényi, A., On the strength of connectedness of a random graph, *Acta Mathematica Scientia Hungary* **12**, 261–267 (1961).

[144] Ergün, G., Human sexual contact network as a bipartite graph, *Physica A* **308**, 483–488 (2002).

[145] Ergün, G. and Rodgers, G. J., Growing random networks with fitness, *Physica A* **303**, 261–272 (2002).

[146] Eriksen, K. A., Simonsen, I., Maslov, S., and Sneppen, K., Modularity and extreme edges of the Internet, Preprint cond-mat/0212001 (2002).

[147] Everitt, B., *Cluster Analysis*, John Wiley, New York (1974).

[148] Faloutsos, M., Faloutsos, P., and Faloutsos, C., On power-law relationships of the internet topology, *Computer Communications Review* **29**, 251–262 (1999).

[149] Fararo, T. J. and Sunshine, M., *A Study of a Biased Friendship Network*, Syracuse University Press, Syracuse (1964).

[150] Farkas, I. J., Derényi, I., Barabási, A.-L., and Vicsek, T., Spectra of "real-world" graphs: Beyond the semicir-

cle law, *Phys. Rev. E* **64**, 026704 (2001).

[151] Farkas, I. J., Derényi, I., Jeong, H., Neda, Z., Oltvai, Z. N., Ravasz, E., Schurbert, A., Barabási, A.-L., and Vicsek, T., Networks in life: Scaling properties and eigenvalue spectra, *Physica A* **314**, 25–34 (2002).

[152] Farkas, I. J., Jeong, H., Vicsek, T., Barabási, A.-L., and Oltvai, Z. N., The topology of the transcription regulatory network in the yeast, Saccharomyces cerevisiae, *Physica A* **381**, 601–612 (2003).

[153] Fell, D. A. and Wagner, A., The small world of metabolism, *Nature Biotechnology* **18**, 1121–1122 (2000).

[154] Ferguson, N. M. and Garnett, G. P., More realistic models of sexually transmitted disease transmission dynamics: Sexual partnership networks, pair models, and moment closure, *Sex. Transm. Dis.* **27**, 600–609 (2000).

[155] Ferrer i Cancho, R., Janssen, C., and Solé, R. V., Topology of technology graphs: Small world patterns in electronic circuits, *Phys. Rev. E* **64**, 046119 (2001).

[156] Ferrer i Cancho, R. and Solé, R. V., Optimization in complex networks, Preprint cond-mat/0111222 (2001).

[157] Ferrer i Cancho, R. and Solé, R. V., The small world of human language, *Proc. R. Soc. London B* **268**, 2261–2265 (2001).

[158] Flake, G. W., Lawrence, S. R., Giles, C. L., and Coetzee, F. M., Self-organization and identification of Web communities, *IEEE Computer* **35**, 66–71 (2002).

[159] Fox, J. J. and Hill, C. C., From topology to dynamics in biochemical networks, *Chaos* **11**, 809–815 (2001).

[160] Frank, O. and Strauss, D., Markov graphs, *J. Amer. Stat. Assoc.* **81**, 832–842 (1986).

[161] Freeman, L., A set of measures of centrality based upon betweenness, *Sociometry* **40**, 35–41 (1977).

[162] Freeman, L. C., Some antecedents of social network analysis, *Connections* **19**, 39–42 (1996).

[163] French, J. R. P., A formal theory of social power, *Psychological Review* **63**, 181–194 (1956).

[164] Fronczak, A., Fronczak, P., and Holyst, J. A., Exact solution for average path length in random graphs, Preprint cond-mat/0212230 (2002).

[165] Fronczak, A., Holyst, J. A., Jedynak, M., and Sienkiewicz, J., Higher order clustering coefficients in Barabasi-Albert networks, *Physica A* **316**, 688–694 (2002).

[166] Gafiychuk, V., Lubashevsky, I., and Stosyk, A., Remarks on scaling properties inherent to the systems with hierarchically organized supplying network, Preprint nlin/0004033 (2000).

[167] Galaskiewicz, J., *Social Organization of an Urban Grants Economy*, Academic Press, New York (1985).

[168] Galaskiewicz, J. and Marsden, P. V., Interorganizational resource networks: Formal patterns of overlap, *Social Science Research* **7**, 89–107 (1978).

[169] Garfield, E., It's a small world after all, *Current Contents* **43**, 5–10 (1979).

[170] Garfinkel, I., Glei, D. A., and McLanahan, S. S., Assortative mating among unmarried parents, *Journal of Population Economics* **15**, 417–432 (2002).

[171] Girvan, M. and Newman, M. E. J., Community structure in social and biological networks, *Proc. Natl. Acad. Sci. USA* **99**, 8271–8276 (2002).

[172] Gleiss, P. M., Stadler, P. F., Wagner, A., and Fell, D. A., Relevant cycles in chemical reaction networks, *Advances in Complex Systems* **4**, 207–226 (2001).

[173] Goh, K.-I., Kahng, B., and Kim, D., Spectra and eigenvectors of scale-free networks, *Phys. Rev. E* **64**, 051903 (2001).

[174] Goh, K.-I., Kahng, B., and Kim, D., Universal behavior of load distribution in scale-free networks, *Phys. Rev. Lett.* **87**, 278701 (2001).

[175] Goh, K.-I., Oh, E., Jeong, H., Kahng, B., and Kim, D., Classification of scale-free networks, *Proc. Natl. Acad. Sci. USA* **99**, 12583–12588 (2002).

[176] Goldberg, D., Nichols, D., Oki, B. M., and Terry, D., Using collaborative filtering to weave an information tapestry, *Comm. ACM* **35**, 61–70 (1992).

[177] Goldwasser, L. and Roughgarden, J., Construction and analysis of a large Caribbean food web, *Ecology* **74**, 1216–1233 (1993).

[178] Goltsev, A. V., Dorogovtsev, S. N., and Mendes, J. F. F., Critical phenomena in networks, *Phys. Rev. E* **67**, 026123 (2003).

[179] Grassberger, P., On the critical behavior of the general epidemic process and dynamical percolation, *Math. Biosci.* **63**, 157–172 (1983).

[180] Grassberger, P., Go with the winners: A general Monte Carlo strategy, *Computer Physics Communications* **147**, 64–70 (2002).

[181] Greenhalgh, D., Optimal control of an epidemic by ring vaccination, *Communications in Statistics: Stochastic Models* **2**, 339–363 (1986).

[182] Grossman, J. W. and Ion, P. D. F., On a portion of the well-known collaboration graph, *Congressus Numerantium* **108**, 129–131 (1995).

[183] Guare, J., *Six Degrees of Separation: A Play*, Vintage, New York (1990).

[184] Guelzim, N., Bottani, S., Bourgine, P., and Kepes, F., Topological and causal structure of the yeast transcriptional regulatory network, *Nature Genetics* **31**, 60–63 (2002).

[185] Guimerà, R., Danon, L., Díaz-Guilera, A., Giralt, F., and Arenas, A., Self-similar community structure in organisations, Preprint cond-mat/0211498 (2002).

[186] Gupta, S., Anderson, R. M., and May, R. M., Networks of sexual contacts: Implications for the pattern of spread of HIV, *AIDS* **3**, 807–817 (1989).

[187] Hammersley, J. M., Percolation processes: II. The connective constant, *Proc. Cambridge Philos. Soc.* **53**, 642–645 (1957).

[188] Harary, F., *Graph Theory*, Perseus, Cambridge, MA (1995).

[189] Hayes, B., Graph theory in practice: Part I, *American Scientist* **88** (1), 9–13 (2000).

[190] Hayes, B., Graph theory in practice: Part II, *American Scientist* **88** (2), 104–109 (2000).

[191] Herrero, C. P., Ising model in small-world networks, *Phys. Rev. E* **65**, 066110 (2002).

[192] Hethcote, H. W., The mathematics of infectious diseases, *SIAM Review* **42**, 599–653 (2000).

[193] Higham, D. J., Greedy pathlengths and small world graphs, Mathematics Research Report 8, University of Strathclyde (2002).

[194] Holland, P. W. and Leinhardt, S., An exponential family of probability distributions for directed graphs, *J. Amer. Stat. Assoc.* **76**, 33–65 (1981).

[195] Holme, P., Edge overload breakdown in evolving networks, *Phys. Rev. E* **66**, 036119 (2002).

[196] Holme, P., Edling, C. R., and Liljeros, F., Struc-

ture and time-evolution of the Internet community pussokram.com, Preprint cond-mat/0210514 (2002).

[197] Holme, P., Huss, M., and Jeong, H., Subnetwork hierarchies of biochemical pathways, Preprint cond-mat/0206292 (2002).

[198] Holme, P. and Kim, B. J., Growing scale-free networks with tunable clustering, *Phys. Rev. E* **65**, 026107 (2002).

[199] Holme, P. and Kim, B. J., Vertex overload breakdown in evolving networks, *Phys. Rev. E* **65**, 066109 (2002).

[200] Holme, P., Kim, B. J., Yoon, C. N., and Han, S. K., Attack vulnerability of complex networks, *Phys. Rev. E* **65**, 056109 (2002).

[201] Hong, H., Choi, M. Y., and Kim, B. J., Synchronization on small-world networks, *Phys. Rev. E* **65**, 026139 (2002).

[202] Hong, H., Kim, B. J., and Choi, M. Y., Comment on "Ising model on a small world network," *Phys. Rev. E* **66**, 018101 (2002).

[203] Huberman, B. A., *The Laws of the Web*, MIT Press, Cambridge, MA (2001).

[204] Huxham, M., Beaney, S., and Raffaelli, D., Do parasites reduce the chances of triangulation in a real food web?, *Oikos* **76**, 284–300 (1996).

[205] Iamnitchi, A., Ripeanu, M., and Foster, I., Locating data in (small-world?) peer-to-peer scientific collaborations, in P. Druschel, F. Kaashoek, and A. Rowstron (eds.), *Proceedings of the First International Workshop on Peer-to-Peer Systems*, no. 2429 in Lecture Notes in Computer Science, pp. 232–241, Springer, Berlin (2002).

[206] Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y., A comprehensive two-hybrid analysis to explore the yeast protein interactome, *Proc. Natl. Acad. Sci. USA* **98**, 4569–4574 (2001).

[207] Jaffe, A. and Trajtenberg, M., *Patents, Citations and Innovations: A Window on the Knowledge Economy*, MIT Press, Cambridge, MA (2002).

[208] Jain, A. K., Murty, M. N., and Flynn, P. J., Data clustering: A review, *ACM Computing Surveys* **31**, 264–323 (1999).

[209] Jain, S. and Krishna, S., Autocatalytic sets and the growth of complexity in an evolutionary model, *Phys. Rev. Lett.* **81**, 5684–5687 (1998).

[210] Jain, S. and Krishna, S., A model for the emergence of cooperation, interdependence, and structure in evolving networks, *Proc. Natl. Acad. Sci. USA* **98**, 543–547 (2001).

[211] Janson, S., Łuczak, T., and Rucinski, A., *Random Graphs*, John Wiley, New York (1999).

[212] Jeong, H., Mason, S., Barabási, A.-L., and Oltvai, Z. N., Lethality and centrality in protein networks, *Nature* **411**, 41–42 (2001).

[213] Jeong, H., Néda, Z., and Barabási, A.-L., Measuring preferential attachment in evolving networks, *Europhys. Lett.* **61**, 567–572 (2003).

[214] Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., and Barabási, A.-L., The large-scale organization of metabolic networks, *Nature* **407**, 651–654 (2000).

[215] Jespersen, S. and Blumen, A., Small-world networks: Links with long-tailed distributions, *Phys. Rev. E* **62**, 6270–6274 (2000).

[216] Jespersen, S., Sokolov, I. M., and Blumen, A., Relaxation properties of small-world networks, *Phys. Rev. E* **62**, 4405–4408 (2000).

[217] Jin, E. M., Girvan, M., and Newman, M. E. J., The structure of growing social networks, *Phys. Rev. E* **64**, 046132 (2001).

[218] Jones, J. H. and Handcock, M. S., An assessment of preferential attachment as a mechanism for human sexual network formation, Preprint, University of Washington (2003).

[219] Jordano, P., Bascompte, J., and Olesen, J. M., Invariant properties in coevolutionary networks of plant-animal interactions, *Ecology Letters* **6**, 69–81 (2003).

[220] Jost, J. and Joy, M. P., Evolving networks with distance preferences, *Phys. Rev. E* **66**, 036126 (2002).

[221] Kalapala, V. K., Sanwalani, V., and Moore, C., The structure of the United States road network, Preprint, University of New Mexico (2003).

[222] Karinthy, F., Chains, in *Everything is Different*, Budapest (1929).

[223] Karoński, M., A review of random graphs, *Journal of Graph Theory* **6**, 349–389 (1982).

[224] Kauffman, S. A., Metabolic stability and epigenesis in randomly connected nets, *J. Theor. Bio.* **22**, 437–467 (1969).

[225] Kauffman, S. A., Gene regulation networks: A theory for their structure and global behavior, in A. Moscana and A. Monroy (eds.), *Current Topics in Developmental Biology 6*, pp. 145–182, Academic Press, New York (1971).

[226] Kauffman, S. A., *The Origins of Order*, Oxford University Press, Oxford (1993).

[227] Kautz, H., Selman, B., and Shah, M., ReferralWeb: Combining social networks and collaborative filtering, *Comm. ACM* **40**, 63–65 (1997).

[228] Keeling, M. J., The effects of local spatial structure on epidemiological invasion, *Proc. R. Soc. London B* **266**, 859–867 (1999).

[229] Kephart, J. O. and White, S. R., Directed-graph epidemiological models of computer viruses, in *Proceedings of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy*, pp. 343–359, IEEE Computer Society, Los Alamitos, CA (1991).

[230] Killworth, P. D. and Bernard, H. R., The reverse small world experiment, *Social Networks* **1**, 159–192 (1978).

[231] Kim, B. J., Trusina, A., Holme, P., Minnhagen, P., Chung, J. S., and Choi, M. Y., Dynamic instabilities induced by asymmetric influence: Prisoners' dilemma game on small-world networks, *Phys. Rev. E* **66**, 021907 (2002).

[232] Kim, B. J., Yoon, C. N., Han, S. K., and Jeong, H., Path finding strategies in scale-free networks, *Phys. Rev. E* **65**, 027103 (2002).

[233] Kim, J., Krapivsky, P. L., Kahng, B., and Redner, S., Infinite-order percolation and giant fluctuations in a protein interaction network, *Phys. Rev. E* **66**, 055101 (2002).

[234] Kinouchi, O., Martinez, A. S., Lima, G. F., Lourenço, G. M., and Risau-Gusman, S., Deterministic walks in random networks: An application to thesaurus graphs, *Physica A* **315**, 665–676 (2002).

[235] Kleczkowski, A. and Grenfell, B. T., Mean-field-type equations for spread of epidemics: The 'small world' model, *Physica A* **274**, 355–360 (1999).

[236] Kleinberg, J. and Lawrence, S., The structure of the Web, *Science* **294**, 1849–1850 (2001).

[237] Kleinberg, J. M., Authoritative sources in a hyperlinked

environment, *J. ACM* **46**, 604–632 (1999).

[238] Kleinberg, J. M., Navigation in a small world, *Nature* **406**, 845 (2000).

[239] Kleinberg, J. M., The small-world phenomenon: An algorithmic perspective, in *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, pp. 163–170, Association of Computing Machinery, New York (2000).

[240] Kleinberg, J. M., Small world phenomena and the dynamics of information, in T. G. Dietterich, S. Becker, and Z. Ghahramani (eds.), *Proceedings of the 2001 Neural Information Processing Systems Conference*, MIT Press, Cambridge, MA (2002).

[241] Kleinberg, J. M., Kumar, S. R., Raghavan, P., Rajagopalan, S., and Tomkins, A., The Web as a graph: Measurements, models and methods, in *Proceedings of the International Conference on Combinatorics and Computing*, no. 1627 in Lecture Notes in Computer Science, pp. 1–18, Springer, Berlin (1999).

[242] Klemm, K. and Eguiluz, V. M., Highly clustered scale-free networks, *Phys. Rev. E* **65**, 036123 (2002).

[243] Klovdahl, A. S., Potterat, J. J., Woodhouse, D. E., Muth, J. B., Muth, S. Q., and Darrow, W. W., Social networks and infectious disease: The Colorado Springs study, *Soc. Sci. Med.* **38**, 79–88 (1994).

[244] Knuth, D. E., *The Stanford GraphBase: A Platform for Combinatorial Computing*, Addison-Wesley, Reading, MA (1993).

[245] Krapivsky, P. L. and Redner, S., Organization of growing random networks, *Phys. Rev. E* **63**, 066123 (2001).

[246] Krapivsky, P. L. and Redner, S., Finiteness and fluctuations in growing networks, *J. Phys. A* **35**, 9517–9534 (2002).

[247] Krapivsky, P. L. and Redner, S., A statistical physics perspective on Web growth, *Computer Networks* **39**, 261–276 (2002).

[248] Krapivsky, P. L. and Redner, S., Rate equation approach for growing networks, in R. Pastor-Satorras and J. Rubi (eds.), *Proceedings of the XVIII Sitges Conference on Statistical Mechanics*, Lecture Notes in Physics, Springer, Berlin (2003).

[249] Krapivsky, P. L., Redner, S., and Leyvraz, F., Connectivity of growing random networks, *Phys. Rev. Lett.* **85**, 4629–4632 (2000).

[250] Krapivsky, P. L., Rodgers, G. J., and Redner, S., Degree distributions of growing networks, *Phys. Rev. Lett.* **86**, 5401–5404 (2001).

[251] Kretzschmar, M., van Duynhoven, Y. T. H. P., and Severijnen, A. J., Modeling prevention strategies for gonorrhea and chlamydia using stochastic network simulations, *Am. J. Epidemiol.* **114**, 306–317 (1996).

[252] Kulkarni, R. V., Almaas, E., and Stroud, D., Evolutionary dynamics in the Bak-Sneppen model on small-world networks, Preprint cond-mat/9908216 (1999).

[253] Kulkarni, R. V., Almaas, E., and Stroud, D., Exact results and scaling properties of small-world networks, *Phys. Rev. E* **61**, 4268–4271 (2000).

[254] Kumar, R., Raghavan, P., Rajagopalan, S., Sivakumar, D., Tomkins, A. S., and Upfal, E., Stochastic models for the Web graph, in *Proceedings of the 42st Annual IEEE Symposium on the Foundations of Computer Science*, pp. 57–65, Institute of Electrical and Electronics Engineers, New York (2000).

[255] Kuperman, M. and Abramson, G., Small world effect in an epidemiological model, *Phys. Rev. Lett.* **86**, 2909–2912 (2001).

[256] Kuperman, M. and Zanette, D. H., Stochastic resonance in a model of opinion formation on small world networks, *Eur. Phys. J. B* **26**, 387–391 (2002).

[257] Lago-Fernández, L. F., Huerta, R., Corbacho, F., and Sigüenza, J. A., Fast response and temporal coherent oscillations in small-world networks, *Phys. Rev. Lett.* **84**, 2758–2761 (2000).

[258] Lahtinen, J., Kertész, J., and Kaski, K., Scaling of random spreading in small world networks, *Phys. Rev. E* **64**, 057105 (2001).

[259] Lahtinen, J., Kertész, J., and Kaski, K., Random spreading phenomena in annealed small world networks, *Physica A* **311**, 571–580 (2002).

[260] Latora, V. and Marchiori, M., Efficient behavior of small-world networks, *Phys. Rev. Lett.* **87**, 198701 (2001).

[261] Latora, V. and Marchiori, M., Economic small-world behavior in weighted networks, Preprint cond-mat/0204089 (2002).

[262] Latora, V. and Marchiori, M., Is the Boston subway a small-world network?, *Physica A* **314**, 109–113 (2002).

[263] Lawrence, S. and Giles, C. L., Accessibility of information on the web, *Nature* **400**, 107–109 (1999).

[264] Leone, M., Vázquez, A., Vespignani, A., and Zecchina, R., Ferromagnetic ordering in graphs with arbitrary degree distribution, *Eur. Phys. J. B* **28**, 191–197 (2002).

[265] Liljeros, F., Edling, C. R., and Amaral, L. A. N., Sexual networks: Implication for the transmission of sexually transmitted infection, *Microbes and Infections* (in press).

[266] Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E., and Åberg, Y., The web of human sexual contacts, *Nature* **411**, 907–908 (2001).

[267] Lloyd, A. L. and May, R. M., How viruses spread among computers and people, *Science* **292**, 1316–1317 (2001).

[268] Łuczak, T., Sparse random graphs with a given degree sequence, in A. M. Frieze and T. Łuczak (eds.), *Proceedings of the Symposium on Random Graphs, Poznań 1989*, pp. 165–182, John Wiley, New York (1992).

[269] Mariolis, P., Interlocking directorates and control of corporations: The theory of bank control, *Social Science Quarterly* **56**, 425–439 (1975).

[270] Maritan, A., Rinaldo, A., Rigon, R., Giacometti, A., and Rodríguez-Iturbe, I., Scaling laws for river networks, *Phys. Rev. E* **53**, 1510–1515 (1996).

[271] Marsden, P. V., Network data and measurement, *Annual Review of Sociology* **16**, 435–463 (1990).

[272] Martinez, N. D., Artifacts or attributes? Effects of resolution on the Little Rock Lake food web, *Ecological Monographs* **61**, 367–392 (1991).

[273] Martinez, N. D., Constant connectance in community food webs, *American Naturalist* **139**, 1208–1218 (1992).

[274] Maslov, S. and Sneppen, K., Specificity and stability in topology of protein networks, *Science* **296**, 910–913 (2002).

[275] Maslov, S., Sneppen, K., and Zaliznyak, A., Pattern detection in complex networks: Correlation profile of the Internet, Preprint cond-mat/0205379 (2002).

[276] May, R. M. and Anderson, R. M., The transmission dynamics of human immunodeficiency virus (HIV), *Philos. Trans. R. Soc. London B* **321**, 565–607 (1988).

[277] May, R. M. and Lloyd, A. L., Infection dynamics on

scale-free networks, *Phys. Rev. E* **64**, 066112 (2001).

[278] Meester, R. and Roy, R., *Continuum Percolation*, Cambridge University Press, Cambridge (1996).

[279] Melin, G. and Persson, O., Studying research collaboration using co-authorships, *Scientometrics* **36**, 363–377 (1996).

[280] Menczer, F. and Belew, R. K., Adaptive retrieval agents: Internalizing local context and scaling up to the Web, *Machine Learning* **39** (2-3), 203–242 (2000).

[281] Menczer, F., Pant, G., Ruiz, M., and Srinivasan, P., Evaluating topic-driven Web crawlers, in D. H. Kraft, W. B. Croft, D. J. Harper, and J. Zobel (eds.), *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 241–249, Association of Computing Machinery, New York (2001).

[282] Merton, R. K., The Matthew effect in science, *Science* **159**, 56–63 (1968).

[283] Milgram, S., The small world problem, *Psychology Today* **2**, 60–67 (1967).

[284] Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U., Network motifs: Simple building blocks of complex networks, *Science* **298**, 824–827 (2002).

[285] Mitchell, M., *Introduction to Genetic Algorithms*, MIT Press, Cambridge, MA (1996).

[286] Mizruchi, M. S., *The American Corporate Network, 1904–1974*, Sage, Beverley Hills (1982).

[287] Molloy, M. and Reed, B., A critical point for random graphs with a given degree sequence, *Random Structures and Algorithms* **6**, 161–179 (1995).

[288] Molloy, M. and Reed, B., The size of the giant component of a random graph with a given degree sequence, *Combinatorics, Probability and Computing* **7**, 295–305 (1998).

[289] Monasson, R., Diffusion, localization and dispersion relations on 'small-world' lattices, *Eur. Phys. J. B* **12**, 555–567 (1999).

[290] Montoya, J. M. and Solé, R. V., Small world patterns in food webs, *J. Theor. Bio.* **214**, 405–412 (2002).

[291] Moody, J., Race, school integration, and friendship segregation in America, *Am. J. Sociol.* **107**, 679–716 (2001).

[292] Moody, J., The structure of a social science collaboration network, Preprint, Department of Sociology, Ohio State University (2003).

[293] Moore, C. and Newman, M. E. J., Epidemics and percolation in small-world networks, *Phys. Rev. E* **61**, 5678–5682 (2000).

[294] Moore, C. and Newman, M. E. J., Exact solution of site and bond percolation on small-world networks, *Phys. Rev. E* **62**, 7059–7064 (2000).

[295] Moreira, A. A., Andrade, Jr., J. S., and Amaral, L. A. N., Extremum statistics in scale-free network models, Preprint cond-mat/0205411 (2002).

[296] Moreno, J. L., *Who Shall Survive?*, Beacon House, Beacon, NY (1934).

[297] Moreno, Y., Gómez, J. B., and Pacheco, A. F., Instability of scale-free networks under node-breaking avalanches, *Europhys. Lett.* **58**, 630–636 (2002).

[298] Moreno, Y., Pastor-Satorras, R., Vázquez, A., and Vespignani, A., Critical load and congestion instabilities in scale-free networks, Preprint cond-mat/0209474 (2002).

[299] Moreno, Y., Pastor-Satorras, R., and Vespignani, A., Epidemic outbreaks in complex heterogeneous networks, *Eur. Phys. J. B* **26**, 521–529 (2002).

[300] Moreno, Y. and Vázquez, A., The Bak-Sneppen model on scale-free networks, *Europhys. Lett.* **57**, 765–771 (2002).

[301] Moreno, Y. and Vázquez, A., Disease spreading in structured scale-free networks, Preprint cond-mat/0210362 (2002).

[302] Morris, M., Data driven network models for the spread of infectious disease, in D. Mollison (ed.), *Epidemic Models: Their Structure and Relation to Data*, pp. 302–322, Cambridge University Press, Cambridge (1995).

[303] Morris, M., Sexual networks and HIV, *AIDS 97: Year in Review* **11**, 209–216 (1997).

[304] Motter, A. E., de Moura, A. P., Lai, Y.-C., and Dasgupta, P., Topology of the conceptual network of language, *Phys. Rev. E* **65**, 065102 (2002).

[305] Motter, A. E. and Lai, Y.-C., Cascade-based attacks on complex networks, *Phys. Rev. E* **66**, 065102 (2002).

[306] Moukarzel, C. F., Spreading and shortest paths in systems with sparse long-range connections, *Phys. Rev. E* **60**, 6263–6266 (1999).

[307] Moukarzel, C. F. and de Menezes, M. A., Shortest paths on systems with power-law distributed long-range connections, *Phys. Rev. E* **65**, 056709 (2002).

[308] Müller, J., Schönfisch, B., and Kirkilionis, M., Ring vaccination, *J. Math. Biol.* **41**, 143–171 (2000).

[309] Newman, M. E. J., Models of the small world, *J. Stat. Phys.* **101**, 819–841 (2000).

[310] Newman, M. E. J., Clustering and preferential attachment in growing networks, *Phys. Rev. E* **64**, 025102 (2001).

[311] Newman, M. E. J., Scientific collaboration networks: I. Network construction and fundamental results, *Phys. Rev. E* **64**, 016131 (2001).

[312] Newman, M. E. J., Scientific collaboration networks: II. Shortest paths, weighted networks, and centrality, *Phys. Rev. E* **64**, 016132 (2001).

[313] Newman, M. E. J., The structure of scientific collaboration networks, *Proc. Natl. Acad. Sci. USA* **98**, 404–409 (2001).

[314] Newman, M. E. J., Assortative mixing in networks, *Phys. Rev. Lett.* **89**, 208701 (2002).

[315] Newman, M. E. J., Spread of epidemic disease on networks, *Phys. Rev. E* **66**, 016128 (2002).

[316] Newman, M. E. J., The structure and function of networks, *Computer Physics Communications* **147**, 40–45 (2002).

[317] Newman, M. E. J., Ego-centered networks and the ripple effect, *Social Networks* **25**, 83–95 (2003).

[318] Newman, M. E. J., Mixing patterns in networks, *Phys. Rev. E* **67**, 026126 (2003).

[319] Newman, M. E. J., Random graphs as models of networks, in S. Bornholdt and H. G. Schuster (eds.), *Handbook of Graphs and Networks*, pp. 35–68, Wiley-VCH, Berlin (2003).

[320] Newman, M. E. J., Barabási, A.-L., and Watts, D. J., *The Structure and Dynamics of Networks*, Princeton University Press, Princeton (2003).

[321] Newman, M. E. J., Forrest, S., and Balthrop, J., Email networks and the spread of computer viruses, *Phys. Rev. E* **66**, 035101 (2002).

[322] Newman, M. E. J., Moore, C., and Watts, D. J., Mean-

field solution of the small-world network model, *Phys. Rev. Lett.* **84**, 3201–3204 (2000).

[323] Newman, M. E. J., Strogatz, S. H., and Watts, D. J., Random graphs with arbitrary degree distributions and their applications, *Phys. Rev. E* **64**, 026118 (2001).

[324] Newman, M. E. J. and Watts, D. J., Renormalization group analysis of the small-world network model, *Phys. Lett. A* **263**, 341–346 (1999).

[325] Newman, M. E. J. and Watts, D. J., Scaling and percolation in the small-world network model, *Phys. Rev. E* **60**, 7332–7342 (1999).

[326] Ozana, M., Incipient spanning cluster on small-world networks, *Europhys. Lett.* **55**, 762–766 (2001).

[327] Padgett, J. F. and Ansell, C. K., Robust action and the rise of the Medici, 1400–1434, *Am. J. Sociol.* **98**, 1259–1319 (1993).

[328] Page, L., Brin, S., Motwani, R., and Winograd, T., The Pagerank citation ranking: Bringing order to the web, Technical report, Stanford University (1998).

[329] Pandit, S. A. and Amritkar, R. E., Random spread on the family of small-world networks, *Phys. Rev. E* **63**, 041104 (2001).

[330] Pastor-Satorras, R. and Rubi, J. (eds.), *Proceedings of the XVIII Sitges Conference on Statistical Mechanics*, Lecture Notes in Physics, Springer, Berlin (2003).

[331] Pastor-Satorras, R., Vázquez, A., and Vespignani, A., Dynamical and correlation properties of the Internet, *Phys. Rev. Lett.* **87**, 258701 (2001).

[332] Pastor-Satorras, R. and Vespignani, A., Epidemic dynamics and endemic states in complex networks, *Phys. Rev. E* **63**, 066117 (2001).

[333] Pastor-Satorras, R. and Vespignani, A., Epidemic spreading in scale-free networks, *Phys. Rev. Lett.* **86**, 3200–3203 (2001).

[334] Pastor-Satorras, R. and Vespignani, A., Epidemic dynamics in finite size scale-free networks, *Phys. Rev. E* **65**, 035108 (2002).

[335] Pastor-Satorras, R. and Vespignani, A., Immunization of complex networks, *Phys. Rev. E* **65**, 036104 (2002).

[336] Pastor-Satorras, R. and Vespignani, A., Epidemics and immunization in scale-free networks, in S. Bornholdt and H. G. Schuster (eds.), *Handbook of Graphs and Networks*, Wiley-VCH, Berlin (2003).

[337] Pękalski, A., Ising model on a small world network, *Phys. Rev. E* **64**, 057104 (2001).

[338] Pennock, D. M., Flake, G. W., Lawrence, S., Glover, E. J., and Giles, C. L., Winners don't take all: Characterizing the competition for links on the web, *Proc. Natl. Acad. Sci. USA* **99**, 5207–5211 (2002).

[339] Pimm, S. L., *Food Webs*, University of Chicago Press, Chicago, 2nd ed. (2002).

[340] Podani, J., Oltvai, Z. N., Jeong, H., Tombor, B., Barabási, A.-L., and Szathmary, E., Comparable system-level organization of Archaea and Eukaryotes, *Nature Genetics* **29**, 54–56 (2001).

[341] Pool, I. de S. and Kochen, M., Contacts and influence, *Social Networks* **1**, 1–48 (1978).

[342] Potterat, J. J., Phillips-Plummer, L., Muth, S. Q., Rothenberg, R. B., Woodhouse, D. E., Maldonado-Long, T. S., Zimmerman, H. P., and Muth, J. B., Risk network structure in the early epidemic phase of HIV transmission in Colorado Springs, *Sexually Transmitted Infections* **78**, i159–i163 (2002).

[343] Price, D. J. de S., Networks of scientific papers, *Science* **149**, 510–515 (1965).

[344] Price, D. J. de S., A general theory of bibliometric and other cumulative advantage processes, *J. Amer. Soc. Inform. Sci.* **27**, 292–306 (1976).

[345] Ramezanpour, A., Karimipour, V., and Mashaghi, A., Generating correlated networks from uncorrelated ones, Preprint cond-mat/0212469 (2002).

[346] Rapoport, A., Contribution to the theory of random and biased nets, *Bulletin of Mathematical Biophysics* **19**, 257–277 (1957).

[347] Rapoport, A., Cycle distribution in random nets, *Bulletin of Mathematical Biophysics* **10**, 145–157 (1968).

[348] Rapoport, A. and Horvath, W. J., A study of a large sociogram, *Behavioral Science* **6**, 279–291 (1961).

[349] Ravasz, E. and Barabási, A.-L., Hierarchical organization in complex networks, *Phys. Rev. E* **67**, 026112 (2003).

[350] Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z., and Barabási, A.-L., Hierarchical organization of modularity in metabolic networks, *Science* **297**, 1551–1555 (2002).

[351] Redner, S., How popular is your paper? An empirical study of the citation distribution, *Eur. Phys. J. B* **4**, 131–134 (1998).

[352] Resnick, P. and Varian, H. R., Recommender systems, *Comm. ACM* **40**, 56–58 (1997).

[353] Rinaldo, A., Rodríguez-Iturbe, I., and Rigon, R., Channel networks, *Annual Review of Earth and Planetary Science* **26**, 289–327 (1998).

[354] Ripeanu, M., Foster, I., and Iamnitchi, A., Mapping the Gnutella network: Properties of large-scale peer-to-peer systems and implications for system design, *IEEE Internet Computing* **6**, 50–57 (2002).

[355] Rodgers, G. J. and Darby-Dowman, K., Properties of a growing random directed network, *Eur. Phys. J. B* **23**, 267–271 (2001).

[356] Rodríguez-Iturbe, I. and Rinaldo, A., *Fractal River Basins: Chance and Self-Organization*, Cambridge University Press, Cambridge (1997).

[357] Roethlisberger, F. J. and Dickson, W. J., *Management and the Worker*, Harvard University Press, Cambridge, MA (1939).

[358] Rothenberg, R., Baldwin, J., Trotter, R., and Muth, S., The risk environment for HIV transmission: Results from the Atlanta and Flagstaff network studies, *Journal of Urban Health* **78**, 419–431 (2001).

[359] Rozenfeld, A. F., Cohen, R., ben Avraham, D., and Havlin, S., Scale-free networks on lattices, *Phys. Rev. Lett.* **89**, 218701 (2002).

[360] Sander, L. M., Warren, C. P., Sokolov, I., Simon, C., and Koopman, J., Percolation on disordered networks as a model for epidemics, *Math. Biosci.* **180**, 293–305 (2002).

[361] Scala, A., Amaral, L. A. N., and Barthélémy, M., Small-world networks and the conformation space of a short lattice polymer chain, *Europhys. Lett.* **55**, 594–600 (2001).

[362] Schwartz, N., Cohen, R., ben-Avraham, D., Barabási, A.-L., and Havlin, S., Percolation in directed scale-free networks, *Phys. Rev. E* **66**, 015104 (2002).

[363] Scott, J., *Social Network Analysis: A Handbook*, Sage Publications, London, 2nd ed. (2000).

[364] Seglen, P. O., The skewness of science, *J. Amer. Soc. Inform. Sci.* **43**, 628–638 (1992).

[365] Sen, P. and Chakrabarti, B. K., Small-world phenomena and the statistics of linear polymers, *J. Phys. A* **34**, 7749–7755 (2001).

[366] Sen, P., Dasgupta, S., Chatterjee, A., Sreeram, P. A., Mukherjee, G., and Manna, S. S., Small-world properties of the Indian railway network, Preprint cond-mat/0208535 (2002).

[367] Shardanand, U. and Maes, P., Social information filtering: Algorithms for automating "word of mouth", in *Proceedings of ACM Conference on Human Factors and Computing Systems*, pp. 210–217, Association of Computing Machinery, New York (1995).

[368] Shen-Orr, S., Milo, R., Mangan, S., and Alon, U., Network motifs in the transcriptional regulation network of Escherichia coli, *Nature Genetics* **31**, 64–68 (2002).

[369] Sigman, M. and Cecchi, G. A., Global organization of the Wordnet lexicon, *Proc. Natl. Acad. Sci. USA* **99**, 1742–1747 (2002).

[370] Simon, H. A., On a class of skew distribution functions, *Biometrika* **42**, 425–440 (1955).

[371] Smith, R. D., Instant messaging as a scale-free network, Preprint cond-mat/0206378 (2002).

[372] Snijders, T. A. B., Markov chain Monte Carlo estimation of exponential random graph models, *Journal of Social Structure* **2** (2) (2002).

[373] Socolar, J. E. S. and Kauffman, S. A., Scaling in ordered and critical random Boolean networks, *PRL* **90**, 068702 (2003).

[374] Söderberg, B., General formalism for inhomogeneous random graphs, *Phys. Rev. E* **66**, 066121 (2002).

[375] Solé, R. V. and Montoya, J. M., Complexity and fragility in ecological networks, *Proc. R. Soc. London B* **268**, 2039–2045 (2001).

[376] Solé, R. V. and Pastor-Satorras, R., Complex networks in genomics and proteomics, in S. Bornholdt and H. G. Schuster (eds.), *Handbook of Graphs and Networks*, pp. 145–167, Wiley-VCH, Berlin (2003).

[377] Solé, R. V., Pastor-Satorras, R., Smith, E., and Kepler, T. B., A model of large-scale proteome evolution, *Advances in Complex Systems* **5**, 43–54 (2002).

[378] Solomonoff, R. and Rapoport, A., Connectivity of random nets, *Bulletin of Mathematical Biophysics* **13**, 107–117 (1951).

[379] Sporns, O., Network analysis, complexity, and brain function, *Complexity* **8** (1), 56–60 (2002).

[380] Sporns, O., Tononi, G., and Edelman, G. M., Theoretical neuroanatomy: Relating anatomical and functional connectivity in graphs and cortical connection matrices, *Cerebral Cortex* **10**, 127–141 (2000).

[381] Stauffer, D., Monte Carlo simulations of Sznajd models, *Journal of Artificial Societies and Social Simulation* **5** (1) (2002).

[382] Stauffer, D., Aharony, A., da Fontoura Costa, L., and Adler, J., Efficient Hopfield pattern recognition on a scale-free neural network, Preprint cond-mat/0212601 (2002).

[383] Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., and Gilles, E. D., Metabolic network structure determines key aspects of functionality and regulation, *Nature* **420**, 190–193 (2002).

[384] Steyvers, M. and Tenenbaum, J. B., The large-scale structure of semantic networks: Statistical analyses and a model for semantic growth, Preprint cond-mat/0110012 (2001).

[385] Strauss, D., On a general class of models for interaction, *SIAM Review* **28**, 513–527 (1986).

[386] Strogatz, S. H., *Nonlinear Dynamics and Chaos*, Addison-Wesley, Reading, MA (1994).

[387] Strogatz, S. H., Exploring complex networks, *Nature* **410**, 268–276 (2001).

[388] Svenson, P., From Néel to NPC: Colouring small worlds, Preprint cs/0107015 (2001).

[389] Szabó, G., Alava, M., and Kertész, J., Structural transitions in scale-free networks, Preprint cond-mat/0208551 (2002).

[390] Sznajd-Weron, K. and Sznajd, J., Opinion evolution in closed community, *Int. J. Mod. Phys. C* **11**, 1157–1165 (2000).

[391] Tadić, B., Dynamics of directed graphs: The World-Wide Web, *Physica A* **293**, 273–284 (2001).

[392] Tadić, B., Temporal fractal structures: Origin of power laws in the World-Wide Web, *Physica A* **314**, 278–283 (2002).

[393] Travers, J. and Milgram, S., An experimental study of the small world problem, *Sociometry* **32**, 425–443 (1969).

[394] Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M., A comprehensive analysis of protein–protein interactions in saccharomyces cerevisiae, *Nature* **403**, 623–627 (2000).

[395] Valverde, S., Cancho, R. F., and Solé, R. V., Scale-free networks from optimal design, *Europhys. Lett.* **60**, 512–517 (2002).

[396] Vázquez, A., Statistics of citation networks, Preprint cond-mat/0105031 (2001).

[397] Vázquez, A., Growing networks with local rules: Preferential attachment, clustering hierarchy and degree correlations, Preprint cond-mat/0211528 (2002).

[398] Vázquez, A., Boguñá, M., Moreno, Y., Pastor-Satorras, R., and Vespignani, A., Topology and correlations in structured scale-free networks, Preprint cond-mat/0209183 (2002).

[399] Vázquez, A., Flammini, A., Maritan, A., and Vespignani, A., Modeling of protein interaction networks, *Complexus* **1**, 38–44 (2003).

[400] Vázquez, A. and Moreno, Y., Resilience to damage of graphs with degree correlations, *Phys. Rev. E* **67**, 015101 (2003).

[401] Vázquez, A., Pastor-Satorras, R., and Vespignani, A., Large-scale topological and dynamical properties of the Internet, *Phys. Rev. E* **65**, 066130 (2002).

[402] Vázquez, A. and Weigt, M., Computational complexity arising from degree correlations in networks, *Phys. Rev. E* **67**, 027101 (2003).

[403] Vazquez, F., Krapivsky, P. L., and Redner, S., Constrained opinion dynamics: Freezing and slow evolution, *J. Phys. A* **36**, L61–L68 (2003).

[404] Wagner, A., The yeast protein interaction network evolves rapidly and contains few redundant duplicate genes, *Mol. Biol. Evol.* **18**, 1283–1292 (2001).

[405] Wagner, A. and Fell, D., The small world inside large metabolic networks, *Proc. R. Soc. London B* **268**, 1803–1810 (2001).

[406] Walsh, T., Search in a small world, in T. Dean (ed.),

Proceedings of the 16th International Joint Conference on Artificial Intelligence, Morgan Kaufmann, San Francisco, CA (1999).

[407] Wang, B.-Y. and Zhang, F., Exact counting of (0,1) matrices with given row and column sums, *Discrete Mathematics* **187**, 211–220 (1998).

[408] Warren, C. P., Sander, L. M., and Sokolov, I., Geography in a scale-free network model, *Phys. Rev. E* **66**, 056105 (2002).

[409] Wasserman, S. and Faust, K., *Social Network Analysis*, Cambridge University Press, Cambridge (1994).

[410] Wasserman, S. and Pattison, P., Logit models and logistic regressions for social networks: I. An introduction to Markov random graphs and p*, *Psychometrika* **61**, 401–426 (1996).

[411] Watts, D. J., Networks, dynamics, and the small world phenomenon, *Am. J. Sociol.* **105**, 493–592 (1999).

[412] Watts, D. J., *Small Worlds*, Princeton University Press, Princeton (1999).

[413] Watts, D. J., A simple model of global cascades on random networks, *Proc. Natl. Acad. Sci. USA* **99**, 5766–5771 (2002).

[414] Watts, D. J., *Six Degrees: The Science of a Connected Age*, Norton, New York (2003).

[415] Watts, D. J., Dodds, P. S., and Newman, M. E. J., Identity and search in social networks, *Science* **296**, 1302–1305 (2002).

[416] Watts, D. J. and Strogatz, S. H., Collective dynamics of 'small-world' networks, *Nature* **393**, 440–442 (1998).

[417] West, G. B., Brown, J. H., and Enquist, B. J., A general model for the origin of allometric scaling laws in biology, *Science* **276**, 122–126 (1997).

[418] West, G. B., Brown, J. H., and Enquist, B. J., A general model for the structure, and allometry of plant vascular systems, *Nature* **400**, 664–667 (1999).

[419] White, H. C., Boorman, S. A., and Breiger, R. L., So-

cial structure from multiple networks: I. Blockmodels of roles and positions, *Am. J. Sociol.* **81**, 730–779 (1976).

[420] White, H. D., Wellman, B., and Nazer, N., Does citation reflect social structure? Longitudinal evidence from the 'Globenet' interdisciplinary research group, Preprint, University of Toronto (2003).

[421] White, J. G., Southgate, E., Thompson, J. N., and Brenner, S., The structure of the nervous system of the nematode C. Elegans, *Phil. Trans. R. Soc. London* **314**, 1–340 (1986).

[422] Wilkinson, D. and Huberman, B. A., A method for finding communities of related genes, Preprint, Stanford University (2002).

[423] Williams, R. J., Berlow, E. L., Dunne, J. A., Barabási, A.-L., and Martinez, N. D., Two degrees of separation in complex food webs, *Proc. Natl. Acad. Sci. USA* **99**, 12913–12916 (2002).

[424] Winfree, A. T., *The Geometry of Biological Time*, Springer, New York, 2nd ed. (2000).

[425] Wormald, N. C., The asymptotic connectivity of labelled regular graphs, *J. Comb. Theory B* **31**, 156–167 (1981).

[426] Young, H. P., The diffusion of innovations in social networks, in L. E. Blume and S. N. Durlauf (eds.), *The Economy as an Evolving Complex System*, vol. 3, Oxford University Press, Oxford (2003).

[427] Zanette, D. H., Critical behavior of propagation on small-world networks, *Phys. Rev. E* **64**, 050901 (2001).

[428] Zekri, N. and Clerc, J. P., Statistical and dynamical study of disease propagation in a small world network, *Phys. Rev. E* **64**, 056115 (2001).

[429] Zhu, J.-Y. and Zhu, H., Introducing small-world network effect to critical dynamics, Preprint cond-mat/0212542 (2002).