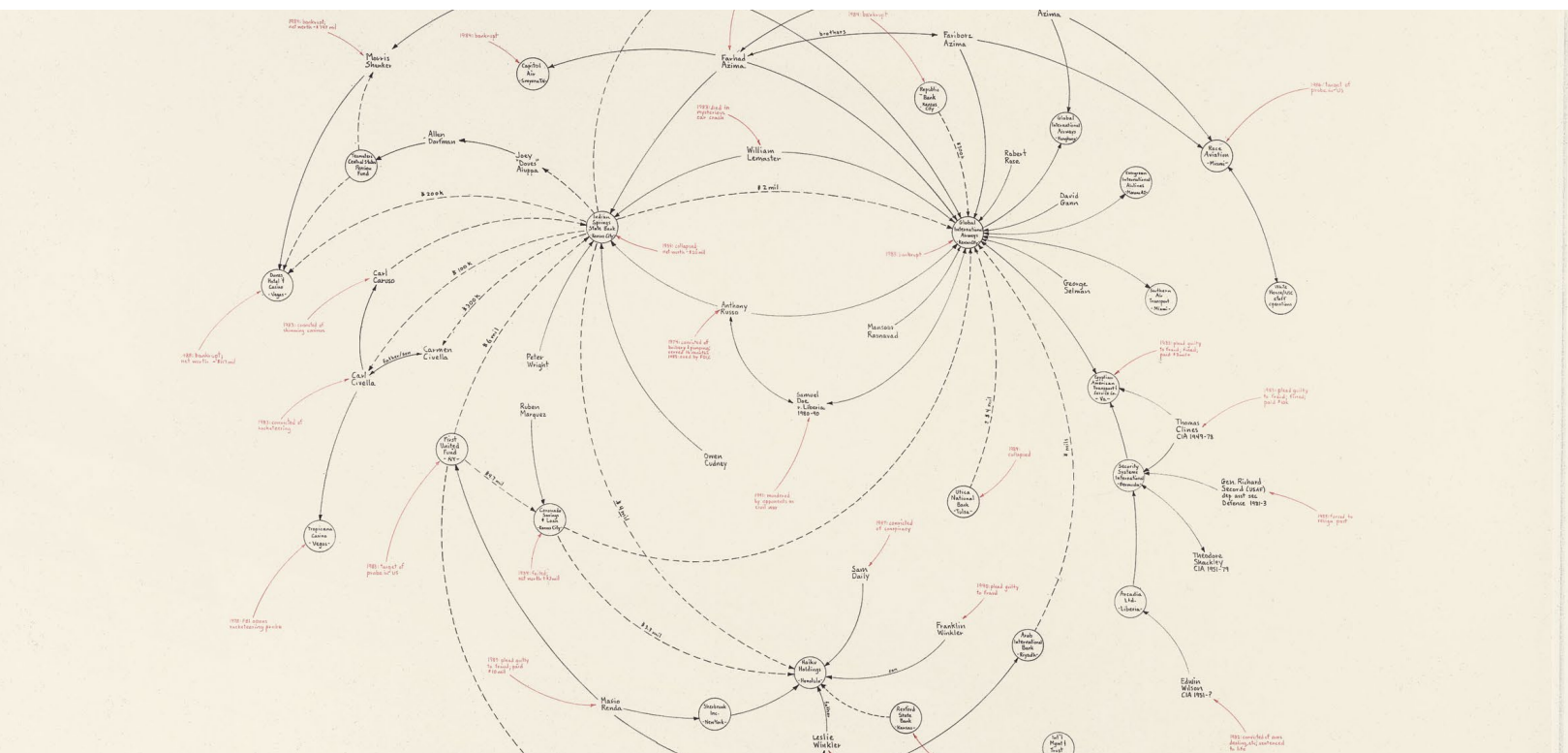


ALBERT-LÁSZLÓ BARABÁSI

NETWORK SCIENCE INTRODUCTION



ACKNOWLEDGEMENTS

MÁRTON PÓSAI
GABRIELE MUSELLA
MAURO MARTINO
ROBERTA SINATRA

PHILIPP HOEVEL
SARAH MORRISON
AMAL HUSSEINI

INDEX

Vulnerability Due to Interconnectivity	1
Networks at the Heart of Complex Systems	2
Two Forces Helped the Emergence of Network Science	3
The Characteristics of Network Science	4
Societal Impact	5
Scientific Impact	6
Summary	7
Homework	8
Bibliography	9

Figure 1.0 (front cover)

Mark Lombardi: Global International Airway and Indian Spring State Bank

Mark Lombardi (1951 – 2000) was an American artist who documented “the uses and abuses of power.” His work was preceded by careful research, resulting in thousands of index cards, whose number began to overwhelm his ability to deal with them. Hence Lombardi began assembling them into hand-drawn diagrams, intended to focus his work. Eventually these diagrams became a form of art on their own [1]. The image shows one such drawing, created between 1977 and 1983 in colored pencil and graphite on paper.



This work is licensed under a
Creative Commons: CC BY-NC-SA 2.0.
PDF V26, 03.09.2014

VULNERABILITY DUE TO INTERCONNECTIVITY

At a first glance the two satellite images of [Figure 1.1](#) are indistinguishable, showing lights shining brightly in highly populated areas and dark spaces that mark vast uninhabited forests and oceans. Yet, upon closer inspection we notice differences: Toronto, Detroit, Cleveland, Columbus and Long Island, bright and shining in (a), have gone dark in (b). This is not a doctored shot from the next Armageddon movie but represents a real image of the US Northeast on August 14, 2003, before and after the blackout that left without power an estimated 45 million people in eight US states and another 10 million in Ontario.

The 2003 blackout is a typical example of a cascading failure. When a network acts as a transportation system, a local failure shifts loads to other nodes. If the extra load is negligible, the system can seamlessly absorb it, and the failure goes unnoticed. If, however, the extra load is too much for the neighboring nodes, they will too tip and redistribute the load to their neighbors. In no time, we are faced with a cascading event, whose magnitude depends on the position and the capacity of the nodes that failed initially.

Cascading failures have been observed in many complex systems. They take place on the Internet, when traffic is rerouted to bypass malfunctioning routers. This routine operation can occasionally create denial of service attacks, which make fully functional routers unavailable by overwhelming them with traffic. We witness cascading events in financial systems, like in 1997, when the International Monetary Fund pressured the central banks of several Pacific nations to limit their credit, which defaulted multiple corporations, eventually resulting in stock market crashes worldwide. The 2009-2011 financial meltdown is often seen as a classic example of a cascading failure, the US credit crisis paralyzing the economy of the globe, leaving behind scores of failed banks, corporations, and even bankrupt states. Cascading failures can be also induced artificially. An example is the worldwide effort to dry up the money supply of terrorist organizations, aimed at crippling their ability to function. Similarly, cancer researchers aim to induce cascading failures in our cells to kill cancer cells.

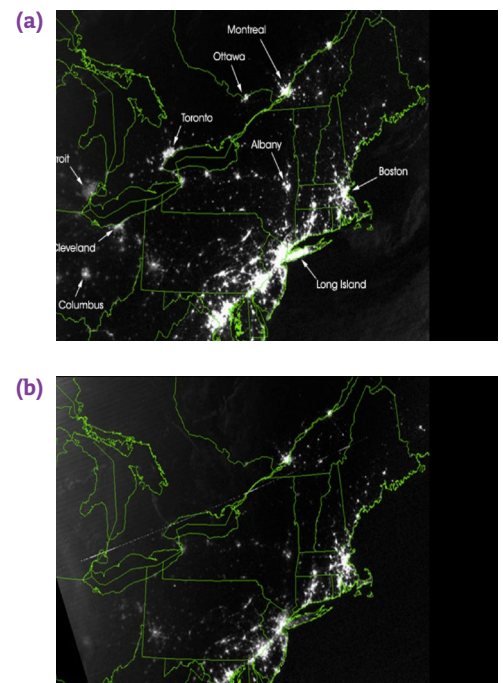


Figure 1.1
2003 North American Blackout

- (a) Satellite image on Northeast United States on August 13th, 2003, at 9:29pm (EDT), 20 hours *before* the 2003 blackout.
- (b) The same as above, but 5 hours *after* the blackout.

The Northeast blackout illustrates several important themes of this book: First, to avoid damaging cascades, we must understand the structure of the network on which the cascade propagates. Second, we must be able to model the dynamical processes taking place on these networks, like the flow of electricity. Finally, we need to uncover how the interplay between the network structure and dynamics affects the robustness of the whole system. Although cascading failures may appear random and unpredictable, they follow reproducible laws that can be quantified and even predicted using the tools of network science.

The blackout also illustrates a bigger theme: *vulnerability due to interconnectivity*. Indeed, in the early years of electric power each city had its own generators and electric network. Electricity cannot be stored, however: Once produced, electricity must be immediately consumed. It made economic sense, therefore, to link neighboring cities up, allowing them to share the extra production and borrow electricity if needed. We owe the low price of electricity today to the power grid, the network that emerged through these pairwise connections, linking all producers and consumers into a single network. It allows cheaply produced power to be instantly transported anywhere. Electricity hence offers a wonderful example of the huge positive impact networks have on our life.

Being part of a network has its catch, however: local failures, like the breaking of a fuse somewhere in Ohio, may not stay local any longer. Their impact can travel along the network's links and affect other nodes, consumers and individuals apparently removed from the original problem. In general interconnectivity induces a remarkable non-locality: It allows information, memes, business practices, power, energy, and viruses to spread on their respective social or technological networks, reaching us, no matter our distance from the source. Hence networks carry both benefits and vulnerabilities. Uncovering the factors that can enhance the spread of traits deemed positive, and limit others that make networks weak or vulnerable, is one of the goals of this book.

NETWORKS AT THE HEART OF COMPLEX SYSTEMS

“I think the next century will be the century of complexity.”

Stephen Hawking

We are surrounded by systems that are hopelessly complicated. Consider for example the society that requires cooperation between billions of individuals, or communications infrastructures that integrate billions of cell phones with computers and satellites. Our ability to reason and comprehend our world requires the coherent activity of billions of neurons in our brain. Our biological existence is rooted in seamless interactions between thousands of genes and metabolites within our cells.

These systems are collectively called *complex systems*, capturing the fact that it is difficult to derive their collective behavior from a knowledge of the system's components. Given the important role complex systems play in our daily life, in science and in economy, their understanding, mathematical description, prediction, and eventually control is one of the major intellectual and scientific challenges of the 21st century.

The emergence of network science at the dawn of the 21st century is a vivid demonstration that science can live up to this challenge. Indeed, behind each complex system there is an intricate network that encodes the interactions between the system's components:

- (a) The network encoding the interactions between genes, proteins, and metabolites integrates these components into live cells. The very existence of this *cellular network* is a prerequisite of life.
- (b) The wiring diagram capturing the connections between neurons, called the *neural network*, holds the key to our understanding of how the brain functions and to our consciousness.

BOX 1.1

COMPLEX

[adj., v. kuh m-pleks, kom-pleks;
n. kom-pleks]

- 1) composed of many interconnected parts; compound; composite: a complex highway system
- 2) characterized by a very complicated or involved arrangement of parts, units, etc.: complex machinery
- 3) so complicated or intricate as to be hard to understand or deal with: a complex problem

Source: *Dictionary.com*

- (c) The sum of all professional, friendship, and family ties, often called the *social network*, is the fabric of the society and determines the spread of knowledge, behavior and resources.
- (d) *Communication networks*, describing which communication devices interact with each other, through wired internet connections or wireless links, are at the heart of the modern communication system.
- (e) The *power grid*, a network of generators and transmission lines, supplies with energy virtually all modern technology.
- (f) *Trade networks* maintain our ability to exchange goods and services, being responsible for the material prosperity that the world has enjoyed since WWII (Figure 1.2).

Networks are also at the heart of some of the most revolutionary technologies of the 21st century, empowering everything from Google to Facebook, CISCO, and Twitter. At the end, networks permeate science, technology, business and nature to a much higher degree than it may be evident upon a casual inspection. Consequently, *we will never understand complex systems unless we develop a deep understanding of the networks behind them.*

The exploding interest in network science during the first decade of the 21st century is rooted in the discovery that despite the obvious diversity of complex systems, the structure and the evolution of the networks behind each system is driven by a common set of fundamental laws and principles. Therefore, notwithstanding the amazing differences in form, size, nature, age, and scope of real networks, most networks are driven by common organizing principles. Once we disregard the nature of the components and the precise nature of the interactions between them, the obtained networks are more similar than different from each other. In the following sections we discuss the forces that have led to the emergence of this new research field and its impact on science, technology, and society.



Figure 1.2
Subtle Networks Behind the Economy

A credit card selected as the 99th object in *The History of the World in 100 Objects* exhibit by the British Museum. This card is a vivid demonstration of the highly interconnected nature of the modern economy, relying on subtle economic and social connections that normally go unnoticed.

The card was issued in the United Arab Emirates in 2009 by the *Hong Kong and Shanghai Banking Corporation*, known as HSBC, a London based bank. The card functions through protocols provided by VISA, a USA based credit association. Yet, the card adheres to Islamic banking principles, which operates in accordance with Fiqhal-Muamalat (Islamic rules of transactions), most notably eliminating interest or *riba*. The card is not limited to muslims in the United Arab Emirates, but is offered in non-Muslim countries as well, to anyone who agrees with its strict ethical guidelines.

TWO FORCES THAT HELPED NETWORK SCIENCE

Network science is a new discipline. One may debate its precise beginning, but by all accounts the field has emerged as a separate discipline only in the 21st century.

Why didn't we have network science two hundred years earlier? After all many of the networks that the field explores are by no means new: metabolic networks date back to the origins of life, with a history of four billion years, and the social network is as old as humanity. Furthermore, many disciplines, from biochemistry to sociology and brain science, have been dealing with their own networks for decades. Graph theory, a prolific subfield of mathematics, has explored graphs since 1735. Is there a reason, therefore, to call network science *the science of the 21st century*?

Something special happened at the dawn of the 21st century that transcended individual research fields and catalyzed the emergence of a new discipline (Figure 1.3). To understand why this happened now and not two hundred years earlier, we need to discuss the two forces that have contributed to the emergence of network science.

THE EMERGENCE OF NETWORK MAPS

To describe the detailed behavior of a system consisting of hundreds to billions of interacting components, we need a map of the system's wiring diagram. In a social system this would require an accurate list of your friends, your friends' friends, and so on. In the WWW this map tells us which webpages link to each other. In the cell the map corresponds to a detailed list of binding interactions and chemical reactions involving genes, proteins, and metabolites.

In the past, we lacked the tools to map these networks. It was equally difficult to keep track of the huge amount of data behind them. The Internet revolution, offering effective and fast data sharing methods and cheap digital storage, fundamentally changed our ability to collect, assemble, share, and analyze data pertaining to real networks.

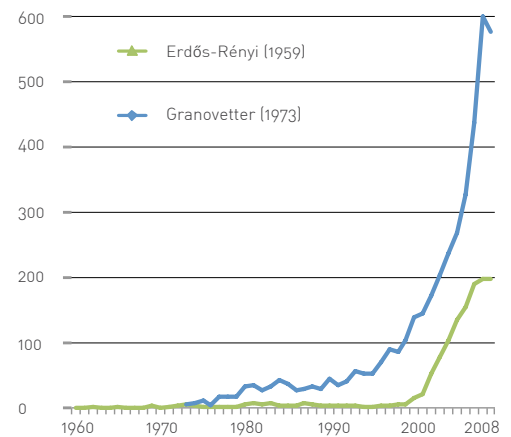


Figure 1.3
The Emergence of Network Science

While the study of networks has a long history, with roots in graph theory and sociology, the modern chapter of network science emerged only during the first decade of the 21st century.

The explosive interest in networks is well documented by the citation pattern of two classic papers, the 1959 paper by Paul Erdős and Alfréd Rényi that marks the beginning of the study of random networks in graph theory [2] and the 1973 paper by Mark Granovetter, the most cited social network paper [3]. The figure shows the yearly citations each paper acquired since their publication. Both papers were highly regarded within their discipline, but had only limited impact outside their field. The explosive growth of citations to these papers in the 21st century is a consequence of the emergence of network science, drawing a new, interdisciplinary attention to these classic publications.

Thanks to these technological advances, at the turn of the millenium we witnessed an explosion of map making (BOX 1.2). Examples range from the CAIDA or DIMES projects that offered the first large-scale maps of the Internet; to the hundreds of millions of dollars spent by biologists to experimentally map out protein-protein interactions in human cells; the efforts made by social network companies, like Facebook, Twitter, or LinkedIn, to develop accurate depositories of our friendships and professional ties; the Connectome project of the US National Institute of Health that aims to systematically trace the neural connections in mammalian brains. The sudden availability of these maps at the end of the 20th century has catalyzed the emergence of network science.

THE UNIVERSALITY OF NETWORK CHARACTERISTICS

It is easy to list the differences between the various networks we encounter in nature or society: the nodes of the metabolic network are tiny molecules and the links are chemical reactions governed by the laws of chemistry and quantum mechanics; the nodes of the WWW are web documents and the links are URLs guaranteed by computer algorithms; the nodes of the social network are individuals and the links represent family, professional, friendship, and acquaintance ties.

The processes that generated these networks also differ greatly: metabolic networks were shaped by billions of years of evolution; the WWW is built by the collective actions of millions of individuals and organizations; social networks are shaped by social norms whose roots go back thousands of years. Given this diversity in size, nature, scope, history, and evolution, one would not be surprised if the networks behind these systems would differ greatly.

A key discovery of network science is that *the architecture of networks emerging in various domains of science, nature, and technology are similar to each other, a consequence of being governed by the same organizing principles. Consequently we can use a common set of mathematical tools to explore these systems.*

This universality is one of the guiding principle of this book: we will not only seek to uncover specific network properties, but each time we ask how widely they apply. We will also aim to understand their origins, uncovering the laws that shape network evolution and their consequences on network behavior.

In summary, while many disciplines have made the important contributions to network science, the emergence of a new field was partly made possible by data availability, offering accurate maps of networks encountered in different disciplines. These diverse maps allowed network scientists to identify the universal properties of various network characteristics. This universality offers the foundation of the new discipline of network science.

BOX 1.2

THE ORIGINS OF NETWORK MAPS

A few of the maps studied today by network scientists were generated with the purpose of studying networks. Most are the byproduct of other projects and morphed into maps only in the hands of network scientists.

- (a) The list of chemical reactions in a cell were discovered one-by-one over a 150 year period by biochemists. In the 1990s they were collected in central databases, offering the first chance to assemble the biochemical networks within a cell.
- (b) The list of actors that play in each movie were traditionally scattered in newspapers, books and encyclopedias. With the advent of the Internet, these data were assembled into central databases, like imdb.com, feeding the curiosity of movie aficionados. The database allowed network scientists to reconstruct the affiliation network behind Hollywood.
- (c) The list of authors of millions of research papers were traditionally scattered in the table of content of thousands of journals. Recently Web of Science, Google Scholar, and other services have assembled them into comprehensive databases, allowing network scientists to reconstruct accurate maps of scientific collaboration networks.

Much of the early history of network science relied on the investigators' ingenuity to recognize and extract networks from preexisting databases. Network science changed that: Today well-funded research collaborations focus on map making, capturing accurate wiring diagrams of biological, communication and social systems.

THE CHARACTERISTICS OF NETWORK SCIENCE

Network science is defined not only by its subject matter, but also by its methodology. In this section we discuss the key characteristics of the approach network science adopted to understand complex systems.

INTERDISCIPLINARY NATURE

Network science offers a language through which different disciplines can seamlessly interact with each other. Indeed, cell biologists, brain scientists (Figure 1.4) and computer scientists alike are faced with the task of characterizing the wiring diagram behind their system, extracting information from incomplete and noisy datasets, and understanding their systems' robustness to failures or attacks.

To be sure, each discipline brings a different set of goals, technical details and challenges, which are important on their own. Yet, the common nature of many issues these fields struggle with has led to a cross-disciplinary fertilization of tools and ideas. For example, the concept of betweenness centrality that emerged in the social network literature in the 1970s, today plays a key role in identifying high traffic nodes on the Internet. Similarly algorithms developed by computer scientists for graph partitioning have found novel applications in identifying disease modules in medicine or detecting communities within large social networks.

EMPIRICAL, DATA DRIVEN NATURE

Several key concepts of network science have their roots in graph theory, a fertile field of mathematics. What distinguishes network science from graph theory is its empirical nature, i.e. its focus on data, function and utility. As we will see in the coming chapters, in network science we are never satisfied with developing abstract mathematical tools to describe a certain network property. Each tool we develop is tested on real data and its value is judged by the insights it offers about a system's properties and behavior.

QUANTITATIVE AND MATHEMATICAL NATURE

To contribute to the development of network science and to properly use its tools, it is essential to master the mathematical formalism behind

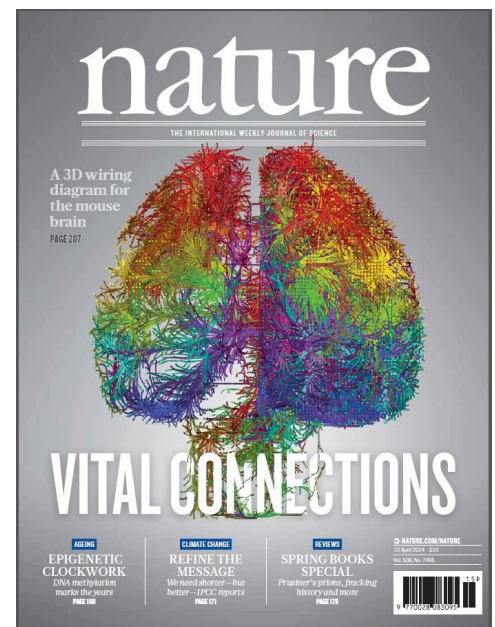


Figure 1.4
Mapping the Brain

An exploding application area for network science is brain research. The wiring diagram of a complete nervous system has long been available for *C. elegans*, a small roundworm, but neuronal connectivity data for larger animals has been missing until recently. That is changing thanks to major efforts by the scientific community to develop technologies that can map out the brain's wiring diagram. The image shows the cover of the April 10, 2014 issue of *Nature*, reporting an extensive map of the laboratory mouse [4] generated by researchers at the Allen Institute in Seattle.

it. Network science borrowed the formalism to deal with graphs from graph theory and the conceptual framework to deal with randomness and seek universal organizing principles from statistical physics. Lately, the field is benefiting from concepts borrowed from engineering, like control and information theory, allowing us to understand the control principles of networks, and from statistics, helping us extract information from incomplete and noisy datasets.

The development of network analysis software has made the tools of network science available to a wider community, even those who may not be familiar with the intellectual foundations and the full mathematical depths of the discipline. Yet, to further the field and to efficiently use its tools, we need to master its theoretical formalism.

COMPUTATIONAL NATURE

Given the size of many of the networks of practical interest, and the exceptional amount of auxiliary data behind them, network scientists are regularly confronted by a series of formidable computational challenges. Hence, the field has a strong computational character, actively borrowing from algorithms, database management and data mining. A series of software tools are available to address these computational problems, enabling practitioners with diverse computational skills to analyze the networks of interest to them.

In summary, a mastery of network science requires familiarity with each of these aspects of the field. It is their combination that offers the multi-faceted tools and perspectives necessary to understand the properties of real networks.

SOCIETAL IMPACT

The impact of a new research field is measured both by its intellectual achievements as well as by its societal impact, indicated by the reach and the potential of its applications. While network science is a young field, its impact is everywhere.

ECONOMIC IMPACT: FROM WEB SEARCH TO SOCIAL NETWORKING

The most successful companies of the 21st century, from Google to Facebook, Twitter, LinkedIn, Cisco, Apple and Akamai, base their technology and business model on networks. Indeed, Google not only runs the biggest network mapping operation that humanity has ever built, generating a comprehensive and constantly updated map of the WWW, but its search technology is deeply interlinked with the network characteristics of the Web.

Networks have gained particular popularity with the emergence of Facebook, the company with the ambition to map out the social network of the whole planet. Facebook was not the first social networking site and it is likely not the last either: An impressive ecosystem of social networking tools, from Twitter to LinkedIn are fighting for the attention of millions of users. Algorithms conceived by network scientists fuel these sites, aiding everything from friend recommendation to advertising.

HEALTH: FROM DRUG DESIGN TO METABOLIC ENGINEERING

Completed in 2001, the human genome project offered the first comprehensive list of all human genes [5, 6]. Yet, to fully understand how our cells function, and the origin of disease, a full list of genes is not sufficient: We also need an accurate map of how genes, proteins, metabolites and other cellular components interact with each other. Indeed, most cellular processes, from food processing to sensing changes in the environment, rely on molecular networks. The breakdown of these networks is responsible for human diseases.

The increasing awareness of the importance of molecular networks

has led to the emergence of *network biology*, a new subfield of biology that aims to understand the behavior of cellular networks. A parallel movement within medicine, called *network medicine*, aims to uncover the role of networks in human disease (Figure 1.5). The importance of these advances is illustrated by the fact that Harvard University in 2012 started the Division of Network Medicine, that employs researchers and medical doctors who apply network-based ideas towards understanding human disease.

Networks play a particularly important role in drug development. The ultimate goal of *network pharmacology* [7] is to develop drugs that can cure diseases without significant side effects. This goal is pursued at many levels, from millions of dollars invested to map out cellular networks, to the development of tools and databases to store, curate, and analyze patient and genetic data.

Several new companies take advantage of the opportunities offered by networks for health and medicine. For example GeneGo collects maps of cellular interactions from the scientific literature and Genomica uses the predictive power behind metabolic networks to identify drug targets in bacteria and humans. Recently major pharmaceutical companies, like Johnson & Johnson, have made significant investments in network medicine, seeing it as the path towards future drugs.

SECURITY: FIGHTING TERRORISM

Terrorism is a malady of the 21st century, requiring significant resources to combat it worldwide. Network thinking is increasingly present in the arsenal of various law enforcement agencies in charge of responding to terrorist activities. It is used to disrupt the financial network of terrorist organizations and to map adversarial networks, helping to uncover the role of their members and their capabilities. While much of the work in this area is classified, several well documented case studies have been made public. Examples include the use of social networks to find Saddam Hussein [10] or those responsible for the March 11, 2004 Madrid train bombings through the examination of the mobile call network. Network concepts have impacted military doctrine as well, leading to the concept of *network-centric warfare*, aimed at fighting low intensity conflicts against terrorist and criminal networks that employ decentralized flexible network organization [11] (Figure 1.6).

Given the numerous potential military applications, it is perhaps not surprising that one of the first academic programs in network science was started at West Point, the US Army Military Academy. Furthermore, starting in 2009 the Army Research Lab devoted over \$300 million to support network science centers across the US.

The knowledge and the capabilities offered by networks can be also abused. Such misuses were well illustrated by the indiscriminate network mapping operation by the National Security Agency [12]. Under the pretext of stopping future terrorist attacks, NSA monitored the



Figure 1.5
Network Biology and Medicine

The cover of two issues of *Nature Reviews Genetics*, the leading review journal in genetics. The journal has devoted exceptional attention to the impact of networks: the 2004 cover focuses on *network biology* [8] (top), the 2011 cover discusses *network medicine* [9] (bottom).

communications of hundreds of millions of individuals, from the US and abroad, rebuilding their social network. With that network scientists have awoken to a new social responsibility: to ensure the ethical use of our tools and knowledge.

EPIDEMICS: FROM FORECASTING TO HALTING DEADLY VIRUSES

While the H1N1 pandemic was not as devastating as it was feared at the beginning of the outbreak in 2009, it gained a special role in the history of epidemics: It was the first pandemic whose course and time evolution was accurately predicted months before the pandemic reached its peak ([Online Resource 1.1](#)) [13]. This was possible thanks to fundamental advances in understanding the role of transportation networks in the spread of viruses.

Before 2000 epidemic modeling was dominated by compartment-based models, assuming that everyone can infect everyone else in the same socio-physical compartment. The emergence of a network-based framework has brought a fundamental change, offering a new level of predictability. Today epidemic prediction is one of the most active applications of network science [13, 14], being used to foresee the spread of influenza or to contain Ebola. It is also the source several fundamental results covered in this book, allowing us to model and predict the spread of biological, digital and social viruses (memes).

The impact of these advances are felt beyond epidemiology. Indeed, in January 2010 network science tools have predicted the conditions necessary for the emergence of viruses spreading through mobile phones [15]. The first major mobile epidemic outbreak that started in the fall of 2010 in China, infecting over 300,000 phones each day, closely followed the predicted scenario.

NEUROSCIENCE: MAPPING THE BRAIN

The human brain, consisting of hundreds of billions of interlinked neurons, is one of the least understood networks from the perspective of network science. The reason is simple: We lack maps telling us which neurons are linked together. The only fully mapped brain available for research is that of the *C. elegans* worm, consisting of only 302 neurons. Detailed maps of mammalian brains could lead to a revolution in brain science, allowing the understanding and curing of numerous neurological and brain diseases. With that brain research could turn it into one of the most prolific application area of network science [16]. Driven by the potential transformative impact of such maps, in 2010 the National Institutes of Health in the U.S. has initiated the *Connectome* project, aimed at developing technologies that could provide accurate neuron-level maps of mammalian brains ([Figure 1.4](#)).

MANAGEMENT: UNCOVERING THE INTERNAL STRUCTURE OF AN ORGANIZATION

While management tends to rely on the official chain of command, it is increasingly evident that the informal network, capturing who really communicates with whom, plays the most important role in the suc-



Online Resource 1.1 Predicting the H1N1 Epidemic

The predicted spread of the H1N1 epidemics during 2009, representing the first successful real-time prediction of a pandemic [13]. The project, relying on data describing the structure and the dynamics of the worldwide transportation network, foresaw that H1N1 will peak out in October 2009, in contrast with the expected January-February peak of influenza. This meant that the vaccines timed for November 2009 were too late, eventually having little impact on the outcome of the epidemic. The success of this project shows the power of network science in facilitating advances in areas of key importance for humanity.

Video courtesy of Alessandro Vespignani.



cess of an organization. Accurate maps of such *organizational networks* can expose the potential lack of interactions between key units, help identify individuals who play an important role in bringing different departments and products together, and help higher management diagnose diverse organizational issues. Furthermore, there is increasing evidence in the management literature that the productivity of an employee is determined by his/her position in this informal organizational network [17].

Therefore, numerous companies, like Maven 7, Activate Networks or Orgnet, offer tools and methodologies to map out the true structure of an organization. These companies offer a host of services, from identifying opinion leaders to reducing employee churn, optimizing knowledge and product diffusion and designing teams with the diversity, size and expertise to be the most effective for specific tasks (Figure 1.8). Established firms, from IBM to SAP, have added social networking capabilities to their business. Overall, network science tools are indispensable in management and business, enhancing productivity and boosting innovation within an organization.

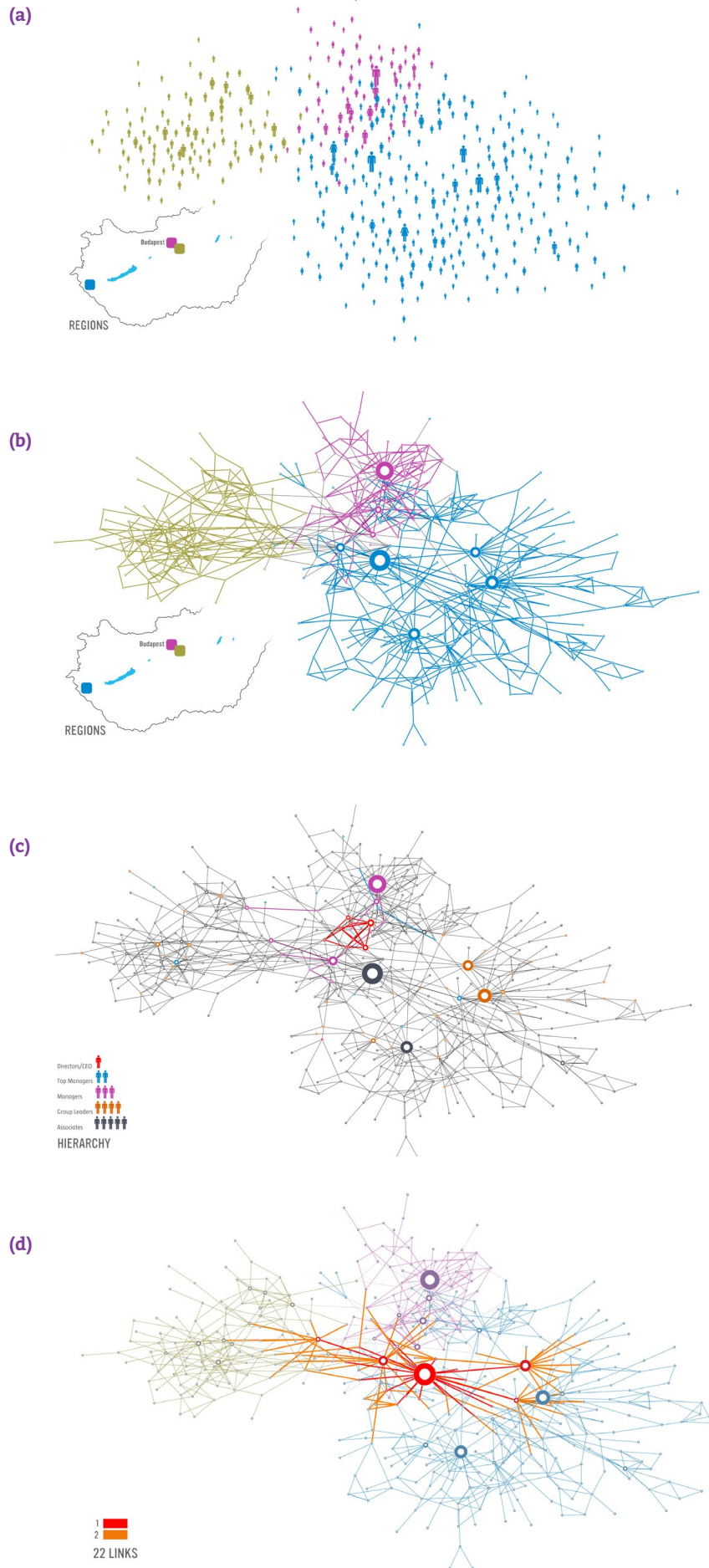


Figure 1.7
Mapping Organizations

(a) Employees of a Hungarian company with three main locations (purple, yellow and blue). The management realized that information reaching the workers about the intentions of the higher management often had nothing to do with their real plans. Seeking to enhance information flow within the company, they turned to Maven 7, a company that applies network science in organizational setting.

(b) Maven 7 developed an online platform to ask each employee to whom do they turn to for advice when it comes to decisions impacting the company. This platform provided the map shown in (b), where two individuals are connected if one nominated the other as his/her source of information on organizational and professional issues. The map identifies several highly influential individuals, appearing as large hubs.

(c) The position of the leadership within the company's informal network, nodes being colored based on their rank within the company. Note that none of the directors, shown in red, are hubs. Nor are the top managers, shown in blue. The hubs come from lower ranks: they are managers, group leaders and associates. The biggest hub, hence the most influential individual, is an ordinary employee, appearing as a gray node in the center.

(d) The links of the largest hub (red) and those two links away from this hub (orange), demonstrate that a significant fraction of employees are at most two links from this hub. But who is this hub? He is the employee in charge of safety and environmental issues. Hence he regularly visits each location and talks with the employees. He is connected to everyone except the top management. With little knowledge of the true intentions of the management, he passes on information that he collects along his trail, effectively running a gossip center.

Should they fire or promote the biggest hub? What is the best solution to this problem?

SCIENTIFIC IMPACT

Nowhere is the impact of network science more evident than in the scientific community. The most prominent scientific journals, from *Nature* to *Science*, *Cell* and *PNAS*, have devoted reviews and editorials addressing the impact of networks on various topics, from biology to social sciences. For example, *Science* has published a special issue on networks, marking the ten-year anniversary of the discovery of scale-free networks [18] (Figure 1.8).

During the past decade each year about a dozen international conferences, workshops, summer and winter schools have focused on network science. A highly successful network science conference series, called Net-Sci, attracts the field's practitioners since 2005. Several general-interest books have made bestseller lists in many countries, bringing network science to the general public. Most major universities offer network science courses, attracting a diverse student body, and in 2014 Northeastern University in Boston and the Central European University in Budapest have launched PhD programs in network science.

To see the impact of networks on the scientific community it is useful to inspect the citation patterns of the most cited papers in the area of complex systems. Each of these papers are citation classics, reporting classic discoveries like the butterfly effect, renormalisation group, spin glasses, fractals and neural networks, and cumulatively amassing anywhere between 2,000 and 5,000 citations. To see how the interest in network science compares to the impact of these foundational papers in Figure 1.9 we compare their citation patterns to the citations of the two most cited network science papers: the 1998 paper on small-world phenomena [19] and the 1999 *Science* paper reporting the discovery of scale-free networks [18]. As one can see, the rapid rise of yearly citations to these two papers is without precedent in the area of complex systems.

Several other metrics indicate that network science is impacting in a defining manner numerous disciplines. For example, in several research fields network papers became the most cited papers in their leading journals:

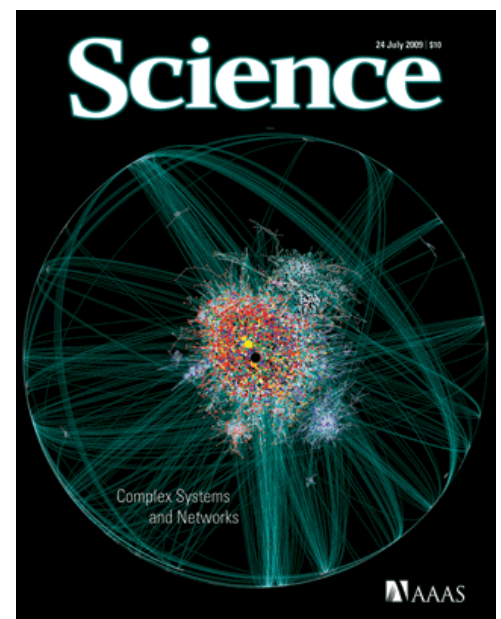


Figure 1.8
Complex Systems and Networks

Special issue of *Science* magazine devoted to networks, published on July 24, 2009, on the 10th anniversary of the 1999 discovery of scale-free networks [18].

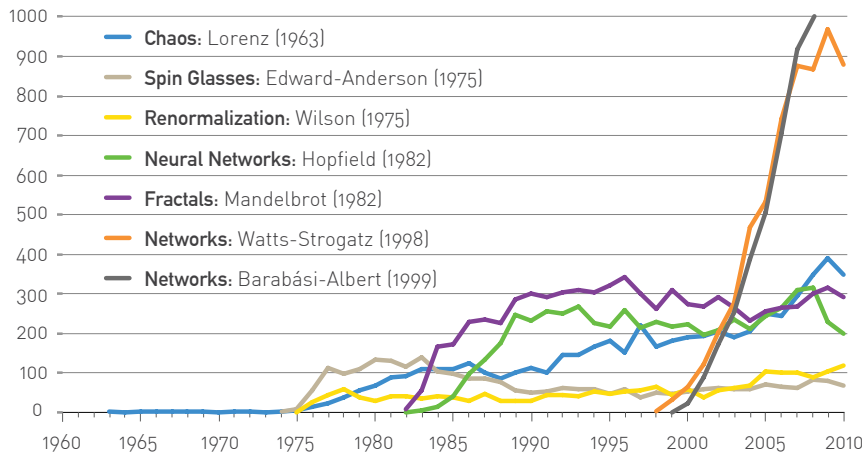


Figure 1.9
Complexity and Network Science

The scientific impact of network science, as seen through citation patterns, compared to the citations of the most cited papers in complexity. The study of complex systems in the 60s and 70s was dominated by Edward Lorenz's 1963 classic work on chaos [20], Kenneth G. Wilson's renormalization group [21], and Samuel F. Edwards and Philip W. Anderson work on spin glasses [22]. In the 1980s the community has shifted its focus to pattern formation, following Benoit Mandelbrot's book on fractals [23] and Thomas Witten and Len Sander's introduction of the diffusion limited aggregation model [24]. Equally influential was John Hopfield's paper on neural networks [25] and Per Bak, Chao Tang and Kurt Wiesenfeld's work on self-organized criticality [26]. These papers continue to define our understanding of complex systems. The figure compares the yearly citations of these landmark papers with the citations of the two most cited papers in network science, the paper by Watts and Strogatz on small world networks and by Barabási and Albert, reporting the discovery of scale-free networks. [18, 19].

- (a) The 1998 paper by Watts and Strogatz in *Nature* on small world phenomena [19] and the 1999 paper by Barabási and Albert in *Science* on scale-free networks [18] were identified by Thompson-Reuters as being among the top ten most cited papers in physical sciences during the decade after their publication. Currently (2011) the Watts-Strogatz paper is the second most cited of all papers published in *Nature* in 1998 and the Barabási-Albert paper is the most cited paper among all papers published in *Science* in 1999.
- (b) Four years after its publication the *SIAM* review by Mark Newman on network science became the most cited paper of any journal published by the Society of Industrial & Applied Mathematics [27].
- (c) *Reviews of Modern Physics*, published since 1929, is the physics journal with the highest impact factor. Until 2012 the most cited paper of the journal was written by Nobel Prize winner Subrahmanyan Chandrasekhar, his classic 1944 review entitled *Stochastic Problems in Physics and Astronomy* [28]. During the 70 years since its publication, the paper gathered over 5,000 citations. Yet, in 2012 it was taken over by the first review of network science published in 2001 entitled *Statistical Mechanics of Complex Networks* [29].
- (d) The paper reporting the discovery that in scale-free networks the epidemic threshold vanishes, by Pastor-Satorras and Vespignani [30], is the most cited paper among the papers published in 2001 by *Physical Review Letters*, shared with a paper on quantum computing.
- (e) The paper by Michelle Girvan and Mark Newman on community discovery in networks [31] is the most cited paper published in 2002 by *Proceedings of the National Academy of Sciences*.
- (f) The 2004 review entitled *Network Biology* [8] is the second most cited paper in the history of *Nature Reviews Genetics*, the top review journal in genetics.

Prompted by this extraordinary enthusiasm within by the scientif-

ic community, network science was examined by the National Research Council (NRC), the arm of the US National Academies in charge of offering policy recommendation to the US government. NRC has assembled two panels, resulting in recommendations summarized in two NRC Reports [32, 33], defining the field of network science (Figure 1.10). These reports not only documented the emergence of a new research field, but highlighted the field's role for science, national competitiveness and security. Following these reports, the National Science Foundation (NSF) in the US established a network science directorate and several Network Science Centers were funded at US universities by the Army Research Labs.

Network science has excited the public as well. This was fueled by the success of several general audience books, like *Linked*, *Nexus*, *Six Degrees* and *Connected* (Figure 1.11). *Connected*, an award-winning documentary by Australian filmmaker Annamaria Talas, has brought the field to our TV screen, being broadcasted all over the world and winning several prestigious prizes (Online Resource 1.2).

Networks have inspired artists as well, leading to a wide range of network-related art projects, and an annual symposium series that brings together artists and network scientists [38]. Fueled by successful movies like *The Social Network* or *Six Degrees of Separation*, and a series of science fiction novels and short stories exploiting the network paradigm, today networks are deeply ingrained in popular culture.

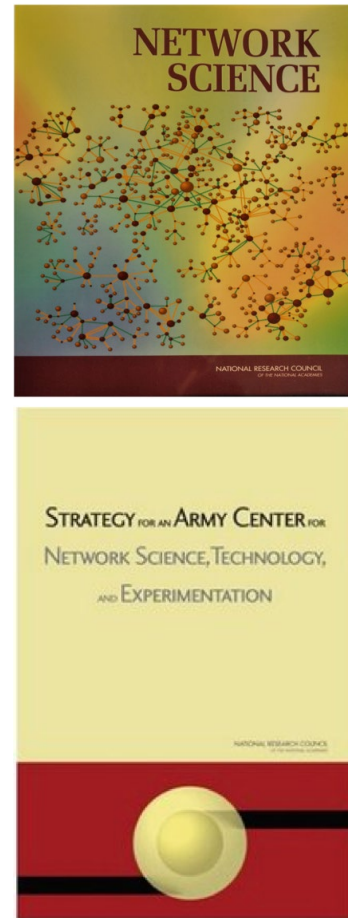
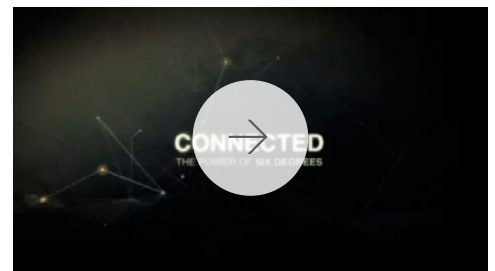


Figure 1.10
National Research Council

Two National Research Council reports on network science have documented the emergence of the new discipline and highlighted its long-term impact on research and national competitiveness [32, 33]. They have recommended dedicated support for the field, prompting the establishment of network science centers at US universities and a network science program within NSF.



Online resource 1.2
Connected

The trailer of the award winning documentary entitled *Connected*, directed by Annamaria Talas, offering an introduction into network science. It features the actor Kevin Bacon and several well-known network scientists.



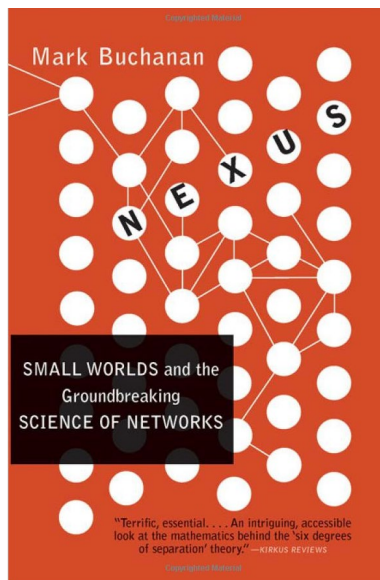
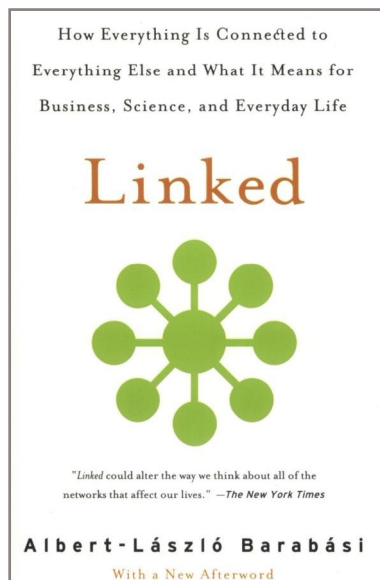
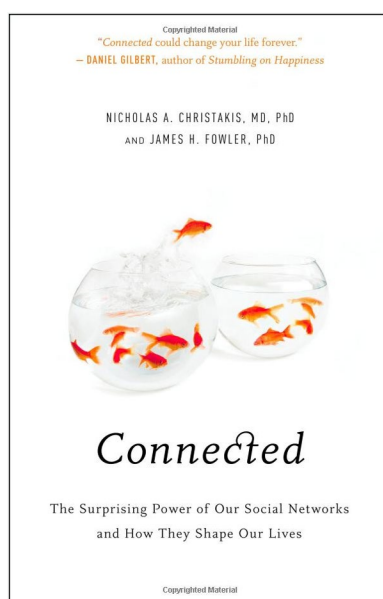
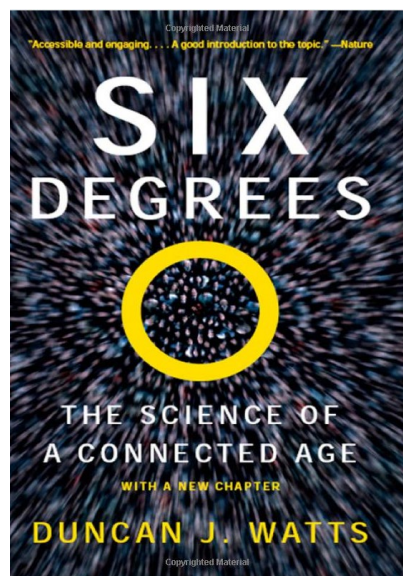


Figure 1.11
Wide Impact

Four widely read books, translated to over twenty languages, have brought network science to the general public [34, 35, 36, 37].



SUMMARY

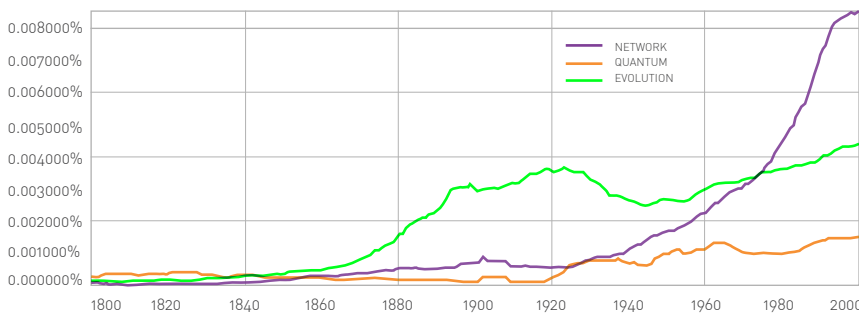


Figure 1.12

The Rise of Networks

The frequency of use of the words *evolution*, *quantum*, and *networks* in books since 1880. The plot indicates the exploding societal awareness of networks in the last decades of the 20th century, laying the ground for the emergence of network science. The plots were generated by Google's *ngram* platform, calculating the fraction of books published in a year that mention *evolution*, *quantum* or *networks*.

While the emergence of network science may appear to have been rather sudden phenomenon (Figures 1.3 & 1.9), the field was responding to a wider social awareness of the role and importance of networks. This is illustrated in Figure 1.12, that shows the usage frequency of words that capture two important scientific revolutions of the past two centuries: *evolution*, the most common term referring to Darwin's theory of evolution, and *quantum*, the most frequently used term when one refers to quantum mechanics. As expected, the use of *evolution* increases after the 1859 publication of Darwin's *On the Origins of Species*. The word *quantum*, first used in 1902, remained virtually absent until the 1920s, when quantum mechanics gained acceptance among physicists and reached public consciousness.

The figure compares these words with the usage of *network*, which enjoyed a spectacular increase following the 1980s, surpassing both *evolution* and *quantum*. While the term *network* has many uses (as do *evolution* and *quantum*), its dramatic rise captures the increasing societal awareness of networks.

There is something common between the advances facilitated by evolutionary theory, quantum mechanics and network science: They are not only important scientific fields with their own intellectual core and body of knowledge, but they are also enabling platforms. Indeed, the current revolution in genetics is built on evolutionary theory and quantum mechanics offers a platform for a wide range of advances in contemporary science, from chemistry to electronics. In a similar fashion, network sci-

ence is an *enabling platform*, offering novel tools and perspectives for a wide range of scientific problems, from social networking to drug design.

Given this exceptional impact networks have both in science and in society, we must master the tools to study and quantify them. The rest of this book is devoted to this worthy subject.

HOMEWORK

1.1. Networks Everywhere

List three different real networks and state the nodes and links for each of them.

1.2. Your Interest

Tell us of the network you are personally most interested in. Address the following questions:

- (a) What are its nodes and links?
- (b) How large is it?
- (c) Can be mapped out?
- (d) Why do you care about it?

1.3. Impact

In your view what would be the area where network science could have the biggest impact in the next decade? Explain your answer.

BIBLIOGRAPHY

[1] J. Richards, R. Hobbs. *Mark Lombardi: Global Networks*. Independent Curators International, New York, 2003.

[2] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae*, 6: 290, 1959.

[3] M. S. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78: 1360, 1973.

[4] S.W. Oh et.al. A mesoscale connectome of the mouse brain. *Nature*, 508: 207-214, 2014.

[5] International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*, 409: 6822, 2001.

[6] J. C. Venter et al. The Sequence of the Human Genome. *Science*, 291: 1304, 2001.

[7] A. L. Hopkins, Network Pharmacology. *Nature Biotechnology*, 25: 1110-1111, 2007.

[8] Z. N. Oltvai and A.-L. Barabási. Network Biology: Understanding the cell's functional organization. *Nature Reviews Genetics*, 5: 101, 2004.

[9] N. Gulbahce, A.-L. Barabási, and J. Loscalzo. Network medicine: A network-based approach to human disease. *Nature Reviews Genetics*, 12: 56, 2011.

[10] C. Wilson. Searching for Saddam: A five-part series on how the US military used social networking to capture the Iraqi dictator. 2010. www.slate.com/id/2245228/.

[11] J. Arquilla and D. Ronfeldt. *Networks and Netwars: The Future of Terror, Crime, and Militancy*. RAND: Santa Monica, CA, 2001.

- [12] A.L. Barabási, Scientists must spearhead ethical use of big data. Politico.com, September 30, 2013.
- [13] D. Balcan, H. Hu, B. Goncalves, P. Bajardi, C. Poletto, J. J. Ramasco, D. Paolotti, N. Perra, M. Tizzoni, W. Van den Broeck, V. Colizza, and A. Vespignani. Seasonal transmission potential and activity peaks of the new influenza A(H1N1): a Monte Carlo likelihood analysis based on human mobility. BMC Medicine, 7: 45, 2009.
- [14] L. Hufnagel, D. Brockmann, and T. Geisel. Forecast and control of epidemics in a globalized world. PNAS, 101: 15124, 2004.
- [15] P. Wang, M. Gonzalez, C. A. Hidalgo, and A.-L. Barabási. Understanding the spreading patterns of mobile phone viruses. Science, 324: 1071, 2009.
- [16] O. Sporns, G. Tononi, and R. Kötter. The Human Connectome: A Structural Description of the Human Brain. PLoS Computational Biology, 1: 4, 2005.
- [17] L. Wu, B. N. Waber, S. Aral, E. Brynjolfsson, and A. Pentland. Mining Face-to-Face Interaction Networks using Sociometric Badges: Predicting Productivity in an IT Configuration Task. Proceedings of the International Conference on Information Systems, Paris, France, December 14-17, 2008.
- [18] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. Science, 286: 509, 1999.
- [19] D. J. Watts and S. H. Strogatz. Collective dynamics of ‘small-world’ networks. Nature, 393: 440, 1998.
- [20] E. N. Lorenz. Deterministic Nonperiodic Flow. Journal of the Atmospheric Sciences, 20: 130, 1963.
- [21] K. G. Wilson. The renormalization group: Critical phenomena and the Kondo problem. Reviews of Modern Physics, 47: 773, 1975.
- [22] S. F. Edwards and P. W. Anderson. Theory of Spin Glasses. Journal of Physics, F 5: 965, 1975.
- [23] B. B. Mandelbrot. *The Fractal Geometry of Nature*. W.H. Freeman and Company. 1982.
- [24] T. Witten, Jr. and L. M. Sander. Diffusion-Limited Aggregation, a Kinetic Critical Phenomenon. Physical Review Letters, 47: 1400, 1981.
- [25] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. PNAS, 79: 2554, 1982.

[26] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality: an explanation of $1/f$ noise. *Physical Review Letters*, 59: 4, 1987.

[27] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*. 45: 167, 2003.

[28] S. Chandrasekhar. Stochastic Problems in Physics and Astronomy. *Reviews Modern Physics*, 15: 1, 1943.

[29] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks. *Reviews Modern Physics*, 74: 47, 2002.

[30] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86: 3200, 2001.

[31] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *PNAS*, 99: 7821, 2002.

[32] National Research Council. *Network Science*. Washington, DC: The National Academies Press, 2005.

[33] National Research Council. *Strategy for an Army Center for Network Science, Technology, and Experimentation*. Washington, DC: The National Academies Press, 2007.

[34] A.-L. Barabási. *Linked: The New Science of Networks*. Perseus Books Group, 2002.

[35] M. Buchanan. *Nexus: Small Worlds and the Groundbreaking Science of Networks*. Norton, 2003.

[36] D. Watts. *Six Degrees: The Science of a Connected Age*. Norton, 2004.

[37] N. Christakis and J. Fowler. *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives*. Back Bay Books, 2011.

[38] M. Schich, R. Malina, and I. Meirelles (Editors). *Arts, Humanities, and Complex Networks [Kindle Edition]*, 2012.