CAMBRIDGE

Cambridge University Press
978-0-521-85740-6 - Complex Social Networks
Fernando Vega-Redondo
Excerpt
More information

CHAPTER 1

Introduction

1.1 THE NEW FIELD OF COMPLEX NETWORKS

In the last decade, the study of complex networks has become a booming field of research with a marked interdisciplinary character. Many different phenomena in the physical, biological, and social worlds can be understood as network based. They build upon some complex (as well as evolving) pattern of bilateral connections among individual entities *and* the overall performance of the system is largely shaped by the intricate architecture of those connections.

A brief review of alternative domains of application should serve to illustrate the rich diversity of phenomena that are distinctly governed by complex networks. This is the task carried out in Subsection 1.1.1, where the primary aim is to illustrate such diversity with empirical illustrations gathered from a large number of different areas. Next, in Subsection 1.1.2 we elaborate on the idea that, given the nature of the endeavour, a genuinely interdisciplinary approach is well in order in the field of complex networks.

1.1.1 Realms of Application and Empirical Evidence

We may start, as the most tangible, with *transportation networks.* These include the connections through which modern economies channel the physical movement of all sorts of commodities and signals. Pertaining, for example, to the conveyance of signals, a paradigmatic instance is of course the internet network, the huge mesh of bilateral connections through which bit-codifying electronic impulses across computers are transferred all around the world. This network has been recently studied by a number of authors, e.g. Govindan and Reddy (1997) [129], Faloutsos *et al.* (1999) [107], Magoni and Pansiot (2001) [194] and Siganos *et al.* (2003) [256].

The exploration of the internet topology can be done at two different levels of detail: at a finer level, focusing on the connections among all routers or, at a coarser one, where only the connections among so-called autonomous systems (or domains) are considered.[1] In either case, the aforementioned studies find a complex network architecture with a broadly skewed distribution of connectivities, i.e. while many nodes (routers or domains) have a few links, some others have many more connections. Or, to be more precise, the empirical observation is that node connectivity is distributed according to a power (or Pareto-like) law. This implies that the degree distribution is *scale-free,* in the sense that the proportional rate ("elasticity") at which the frequency decays with higher degrees is the *same at all scales*. Such an absence of a well-defined scale is often regarded as one of the key factors explaining the intricate complexity of the internet network. In a sense, it is an indication that all scales of the phenomenon are relevant and thus cannot be ignored.

In addition to studying the "physical" internet network, there have also been substantial efforts devoted to understanding the essential features of the "virtual" network defined by the World Wide Web (WWW). In this case, nodes represent the different webpages and a (directed) link joining a webpage to some other is taken to exist if a hyperlink to the latter is found in the former – see e.g. Albert *et al.* (1999) [5] and Kleinberg *et al.* (1999) [176]. The WWW is what might be called an *informational network* since the links present at any given webpage have informational content rather than represent a physical connection.

Another phenomenon whose informational flows can be usefully represented through networks is that of citation, either in patents (cf. Jaffe and Trajtenberg (1996) [163]) or scientific papers (cf. Otte and Rousseau (2002) [229]). In the latter case, for example, the nodes represent scientific papers and a link (again directed) is taken to exist from some paper A to another paper B if A cites B. An early precursor in the study of such citation networks was Price (1965) [239], while a more recent study has been undertaken by Redner (1998) [247].

Interestingly enough, the distribution of connectivities displayed by informational networks so disparate as those of scientific citation and the WWW also happen to be scale-free (at least if we focus on the number of in-connections *received* by each paper or webpage). Moreover, the frequency decay for higher-connectivity nodes (as reflected by the exponent of the power distribution) is so slow that the entailed dispersion of node connectivities is very high – so high, in fact, that second-order moments diverge. This means that the average connectivity cannot be conceived as "typical" and, therefore, there is not even a meaningful way of speaking of a characteristic scale for the connectivity of the network.[2]

[1] While routers are specialized computers directing the flow of internet traffic, the so-called autonomous systems are subnetworks composed of routers that are all under the same administrative control. Two such autonomous systems are understood to be linked if there is at least a pair of routers, one in each system, that are connected.

[2] To make sure, notice the distinction between the notion that a distribution is *scale-free* (which means that is defined by a power law) and the statement that the induced node connectivity *lacks a characteristic scale* (usually identified with infinite, or very high, second-order moments). Even though they are sometimes equated (in part, due to the similar terminology used in both cases),

An additional striking feature that has been documented about the WWW concerns the distance between its nodes, as measured by how many (hyper-)links separate them. Despite the many millions of nodes forming part of the WWW, their average distance turns out to be exceedingly low. For instance, Broder *et al.* (2000) [43] sample (through "crawling strategies") a huge subset of the WWW including more than 200 million nodes and find that their average distance is around 16. This means that, on average, an arbitrary pair of web-pages in the sample are only 16 "clicks" away, following a chain of hyperlinks connecting them.

The networks such as the WWW that enjoy short average distances between their nodes are often said to satisfy the *small-world property* (see Watts (1999) [284] and Buchanan (2002) [44]). This important property has been found to hold in a wide array of large real-world networks, even if these networks display substantial differences pertaining to many other topological features (see Amaral *et al.* (2000) [10]). In this respect, it is worth stressing that the small-world property does *not* require (or implicitly presume) either a broad dispersion in connectivities or that these be distributed in a scale-free fashion.

Indeed, neither broad nor scale-free distributions can be expected to arise in a number of interesting cases where, nevertheless, the average distance between nodes is known to be quite short. This applies, for instance, to contexts where the nature of the phenomenon at hand is such that establishing new links is rather difficult; or it is prohibitively costly to maintain them beyond a certain number; or they tend to age and then vanish at a relatively fast rate. To illustrate the point, we may return to the realm of transportation networks and refer to the data gathered on airline networks by Guimerà *et al.* (2003*b*) [146], or the evidence on networks of electric power distribution reported by Watts and Strogatz (1998) [288]. These networks satisfy the small-world property. And in both of them, the aforementioned circumstances on linking costs apply as well quite naturally.[3] This seems to explain why neither of them is found to display a scale-free distribution of connectivities, in contrast with the evidence outlined above for other transportation and informational networks. But, clearly, analogous considerations may apply as starkly to other kind of networks, such as those arising in biological and social contexts. Next, we review some interesting examples in both of these domains.

Biology has been a fertile area of network applications in recent years. One first (and vast) subarea of research has been molecular biology, the aim being to understand various molecular processes such as metabolic reactions (Jeong *et al.* (2000) [166] and Fell and Wagner (2000) [110]), gene regulation (Kauffman (1993) [170] and Jeong *et al.* (2001) [165]), or the folding of proteins and other polymers (Scala *et al.* (2001) [254] and Amaral *et al.* (2000) [10]).

the former notion implies the latter only if the decay displayed by the degree distribution is not too steep. See Sections 2.1 and 2.5 for an elaboration on these matters.

[3] For example, congestion alone makes adding a new connection to an already busy airport very costly.

In the latter case, for example, each feasible conformation of the polymer chain is identified with a node in the network and a link is defined between two nodes if there is a conformational change of a local character (i.e. one that affects few monomers) that permits switching from one to the other. In a simple two-dimensional representation of the problem, Scala *et al.* (2001) [254] show that the resulting network not only displays the small-world property and a characteristic connectivity but also exhibits significant clustering – i.e. there is a high probability that any two nodes that are neighbors of a third node be neighbors themselves (see Section 2.1 for a formal definition).

High clustering is an indication of a marked local structure (or, analogously, low local dimensionality). It will be seen to have multifarious and important implications for the analysis of networks – in particular for social networks, as explained below. In much of the network literature, it is common to label simply as a *small world* any network that satisfies the twin requirements of short distances (i.e. what has been called the small-world property) *and* high clustering. We shall also abide by this terminology throughout.[4]

Other far-reaching applications of networks in the field of biology have centered on the study of neural networks (Koch and Laurent (1999) [179]) and food webs (Williams and Martinez (2000) [291]). Let us take up each of them in turn. In a sense, we may regard neural networks as a kind of "transportation networks," fulfilling for living beings a role analogous to the internet network in a modern society – i.e. the conveyance of signals (now of an electrochemical nature) across neurons. The best known research that has succeeded in mapping *completely* the neural network of a living being was undertaken by White *et al.* (1986) [290] for the nematode worm *C. elegans*. Abstracting from neurological detail, Watts and Strogatz (1998) [288] translated that mapping into a corresponding (undirected) network of neurons. Interestingly, they found that they are small worlds, i.e. display short distances and a significant degree of clustering.

Food webs, on the other hand, can also be conceived as networks, the nodes standing for the different species in a certain ecological environment and the links embodying predator–prey relationships. Pursuing this approach, specific food webs have been constructed for many disparate ecological systems. To focus on just one example, consider the evidence reported by Montoya and Solé (2002) [205] for three different large ecosystems, two aquatic and one terrestrial. In all three cases, these authors encounter a complex small-world topology with short distances and high clustering.[5] It is remarkable, however, that unlike what occurs in many other small-world networks, the distributions

[4] The expression "small world" was used in the seminal paper of Watts and Strogatz (1998) [288] that introduced a very stylized formalization of the notion. Their model is discussed at some length in Section 2.4.

[5] Naturally, a food web is inherently directed in view of the asymmetry of the predator–prey relationship. This asymmetry notwithstanding, empirical studies in this area (including Montoya and Solé (2002) [205] itself) often abstract from this feature and work with an undirected network representation.

of connectivities is scale-free with a slow decay. This points to environmental conditions that are rich enough to allow at least some species to enjoy substantial trophic versatility.

To conclude our review, let us now turn to the context that is the primary object of this monograph: *social networks*. In the social sciences, the earliest efforts to understand the pattern of inter-agent relationships through "sociograms" – a set of points representing agents and edges joining some of them – goes back to the early work of Moreno (1934) [206], a European socio-psychologist who emigrated in the 1920's to the USA, where he founded the journal *Sociometry*. A mathematical formalization of his ideas, which relies on the concepts and tools of Graph Theory, was later undertaken by Cartwright and Harary (1956) [57]. They explored, in particular, the implications of the innovative "equilibrium-like" notion of structural balance, originally proposed by Heider (1946, 1958) [151, 152]. Heuristically, a network structure is said to be balanced if it equilibrates the overall "tensions" induced by agents' behavior and attitudes toward others (positive and negative).

With such a novel network perspective, lively groups of scholars (psychologists, anthropologists, and sociologists) arose at both sides of the Atlantic, undertaking both theoretical and empirical research. In Europe, the leading group was based at Manchester University – cf. Barnes (1954) [25], Bott (1957) [41], and Mitchell (1969) [201]. In the United States, it was centered at Harvard University around Harrison White (cf. White *et al.* (1976) [290]), followed by Nancy Lee (1969) [188] and Mark Granovetter (1973) [136]. A very useful account of these early developments in social network analysis can be found in Scott (2000, Ch. 2) [255]. This handbook also contains a review of the modern tools and applications in the field of social networks, for which a more complete account can be found in the encyclopedic monograph of Wasserman and Faust (1994) [280].

Unfortunately, the real-world networks that have been studied in detail by the sociological literature mostly focus on small setups, which cannot possibly display the overall complexity that is our main concern here. By way of illustration, we may refer to the empirical cases discussed by Wasserman and Faust (1994, Ch. 2) [280]. They range from the 21 managers of a small firm (Krackhardt (1987) [181]), the 16 leading families of 15th century Florence (Padgett and Ansell (1993) [230]), the 50 researchers of a scientific conference (Freeman (1984) [114]), or the 26 CEO of large firms headquartered in the Minneapolis/St. Paul metropolitan area (Galaskiewicz (1985) [117]). Even the empirical studies that were largely motivated as applications of Rapoport's (1957) [244] *theory of random and biased nets* (an early precursor of the modern theory of complex networks – see Chapter 2) have considered only relatively small contexts. A paradigmatic example is the study conducted by Fararo and Sunshine (1964) [109] – see also Rapoport and Horvath (1961) [245] – who used the individual responses obtained from 417 students of a high school at Madison (Wisconsin) to construct the "sociogram" (network) representing their pattern of friendship.

Only recently, with the wealth of information (and the ability to process it) afforded by modern information technologies, has it become possible to gather data on real-world social networks in a truly large scale.[6] The range of empirical evidence available is still quite limited, but two particular contexts have received special attention: research collaboration and email communication.

Research collaboration is a widespread phenomenon, both among academic scientists and industry researchers. In the academic realm, recent empirical work on collaboration networks has covered a wide variety of disciplines. For example, Newman (2001) [212] has studied the fields of physics, biomedical research, and computer science; Grossman (2002) [143] has focused on mathematics; and Goyal *et al.* (2003) [132] have considered academic economists. The number of authors involved in each case is quite large. It ranges from 11 994 in computer sciences to 52 909 in physics, 81 217 in economics, 337 454 in mathematics, and 1 388 989 in biomedical research, all pertaining to the publication window 1995–99. Qualitatively, the main regularities observed can be summarized as follows.

- Firstly, all of the aforementioned studies report short average distances, which never exceed 10 in any of them.
- A second important observation is that clustering is high. Specifically, the probability that two coauthors of any given researcher be coauthors themselves is above 0.4 in physics and computer science, although significantly lower in the other disciplines (the lowest, 0.072, occurs in biomedical research). Clustering, therefore, must be regarded as quite high in every case, since the populations involved are very large.
- Thirdly, concerning the distribution of connectivities (i.e. of the number of coauthors), one common observation is that, among those academics lying in the lower and middle range, it is scale-free. In contrast, when we move to the higher range, a truncation appears with the frequency of researchers that display a larger number of coauthors decaying much more steeply. Naturally, the latter feature is to be expected in the present context, given the sharp limitations in time and effort faced by authors in writing scientific papers.

Overall, we may confidently conclude that research collaboration networks, at least in the above considered scientific disciplines, are small worlds. Incidentally, it is interesting to note that similar conclusions have also been

[6] In fact, this same ability to handle large data sets effectively has been used by the recent empirical literature in sociology to obtain extensive (and therefore statistically significant) data on *small-* or *mid-size* networks through a large number of *independent materializations* observed in separate groups. Consider, for example, the work of Lubbers (2003) [193] who integrates data on 20,000 students divided in 800 classes to estimate the key considerations affecting class network structure. She relies on the exponential random graph model (also known as the $p*$ model) originally proposed by Frank and Strauss (1986) [113] – see also Wasserman and Pattison (1996) [281]. Another interesting recent example of such a *multilevel* approach is the study by Snijders and Baerveldt (2003) [258] on the spread of delinquent behavior, again integrating behavior (at a “macro” level) of a number of different school classes (at the “micro” level).

obtained concerning collaboration in *industrial* research in a number of different contexts. A good case in point is the network of Italian inventors studied by Balconi *et al.* (2002) [22], where two inventors are taken to be connected if they have collaborated in a joint patent. These networks also happen to display a small-world topology, across a wide range of different industries (chemicals, electronics, consumer goods, etc.).

As mentioned, another social context where large-scale empirical evidence has become recently available is electronic (email) communication.[7] Specifically, Ebel *et al.* (2002) [93] and Guimerà *et al.* (2003*a*) [144] study two different networks representing email communication in university environments. Both of them find small-world features (i.e. low distances and high clustering) in the corresponding network. But whereas Ebel *et al.* (2002) [93] report a broad distribution of connectivities (more specifically, scale-free with a large exponential cut-off), Guimerà *et al.* (2003*a*) [144] identify an exponential distribution with fast decay (and thus a well-defined characteristic scale). This disparity seems to be grounded on the fact that the latter paper chooses to discard bulk email, under the belief that it does not reflect genuine communication. A quite related piece of empirical evidence has been studied by Newman *et al.* (2002) [219], who focus on address books of a large university system. Here, the idea is that individuals keep the email addresses of only those with whom they are in frequent email contact. As by-now expected, they obtain a network with small-world topology. It is also interesting to observe that its connectivity distribution is exponential. This is in line with the conclusion obtained by Guimerà *et al.* (2003*a*) [144], suggesting that when effective communication is involved, individuals face a binding limit on their possible connectivity. From a general viewpoint, this can be viewed as a further illustration of a formerly advanced tenet: a network (be it social or biological) should display a well-defined characteristic connectivity whenever the establishment or maintenance of links is a costly affair.

Lastly, an additional important observation that vertebrates heuristically much of the empirical evidence available on large social networks is that individuals typically do not find it difficult to reach each other. This was the original meaning of the term "small world," coined by Stanley Milgram (1967) [200] to describe a situation where, despite facing a very large population, an individual can typically succeed in *contacting* any other in the population through a short path of intermediaries.[8] Specifically, one of the experiments reported in that seminal article involved delivering a letter from randomly selected individuals

[7] See also the empirical evidence on long-distance phone calls collected by Aiello *et al.* (2000) [3].

[8] As explained in a subsequent article of Stanley Milgram with his student Jeffrey Travers (cf. Travers and Milgram (1969) [270]), the problem was motivated in part as a test of whether some of the related insights obtained by Rapoport and coworkers (cf. in particular the aforementioned Rapoport and Horvath (1961) [245]) extend to a genuinely large and heterogenous population. In fact, the problem was first posed in an even earlier unpublished paper by Ithiel de Sola Pool and Manfred Kochen, which was later published as the inaugural paper of the journal Social Networks, de Sola Pool and Kochen (1978) [260].

living in Nebraska to a stockbroker working in Massachusetts. There was a single simple rule to be respected by the participants: the letter should be sent to someone whom the sender knew on a first-name basis. In the end, a significant fraction of letters arrived to their destination, requiring a median number of steps around six. This, of course, is a direct confirmation that a short path joining them indeed existed in the social network. But, as stressed by some of the latest papers reconsidering the issue (see e.g. Kleinberg (2000 *a,b*) [173, 174]), the truly remarkable feature about the world being "small" is not so much that short paths exist between heterogenous individuals but that these paths can indeed be found! The intriguing aspect of the problem, therefore, concerns the *searchability* – rather than the mere existence – of short paths.

The empirical side of this phenomenon has been recently revisited by Dodds *et al.* (2003) [83], who set up a large-scale, internet-based search experiment similar to that of Milgram's. Eighteen heterogenous targets were selected around the world and more than 24 000 individuals were involved in initiating the search chain. Unfortunately, only a scant 384 chains were completed by reaching the destination, which casts some doubts on the interpretation of the exercise. In particular, as the authors themselves suggest, it underscores the importance of individual incentives (besides searchability) as a primary factor in chain completion. However, when suitably accounting for a reasonable rate of attrition (that is assumed random), they find that Milgram's original conclusions are somewhat reconfirmed: the median number of steps can be estimated to range between 5 and 7.

1.1.2 An Interdisciplinary Approach

The profusion of examples just reviewed suggests the idea that a truly interdisciplinary approach to the study of complex networks may be quite fruitful. A common area of research thus arises where physics, biology, and the social sciences share not only objectives and insights, but also organizing concepts and analytical tools. More specifically, we can single out the following three levels of convergence among these different disciplines.

First, we find that complex networks not only emerge in quite different contexts (from molecular biology to economics, or from ecological systems to the World Wide Web) but also tend to display some analogous topological features. For example, we have seen that large networks with short distances and high clustering (what has been labelled a "small world") appear to prevail in many such contexts. It is conceivable, therefore, that it could respond to similar mechanisms of network formation.

Clustering, specifically, might often be the result of intense *local* search (appropriately defined in the context at hand) in the creation of new links. This introduces a force toward transitivity into link formation that, overall, must tend to enhance clustering. Concerning short typical distances, on the other hand, they could simply follow from the operation of some *global* mechanism that establishes just a few *long-range* links acting as effective shortcuts. Finally, another dimension where general principles might also be at work pertains to the distribution of node connectivities. As hinted before, it may well be that whether or not node connectivity displays a characteristic scale simply depends on the magnitude of linking costs, whether these are biological, ecological, or economic in nature. In general, therefore, we find that striving to unveil common mechanisms and principles that may apply to a wide range of different setups should be one of the distinctive marks of the modern field of complex networks.

A second related basis of interdisciplinary convergence involves the nature of the questions being asked. Thus, for example, some common questions that naturally arise in complex networks, be they biological, technological, or social concern the following phenomena.

- *Robustness*, i.e. the resilience of certain network features (e.g. its overall connectivity, which may crucially affect performance) to the operation of occasional perturbations.
  In this respect, a natural question to ask is how such robustness is affected by the topological characteristics of the network. Relatedly, a further interesting question to pose is when and why may it matter that the perturbations be random ("errors") or guided ("attack"). See e.g. Albert *et al.* (2000) [6], Solé and Montoya (2001) [261], and Carlson and Doyle (2002) [54] for a discussion of these matters in different contexts and with a different perspective.
- *Search*, i.e. the procedure by which individual nodes may look for, and then access, disperse information.
  Here, some natural questions are the following. Is the effectiveness of search influenced by network topology? Are different search algorithms better suited to alternative topologies? Can one design the network architecture so that search is optimized? See e.g. Adamic *et al.* (2001, 2003) [2, 1] and Kleinberg (2000 *a,b*) [173, 174].
- *Diffusion*, i.e. the multifaceted phenomenon that governs the gradual spread over time of any kind of signals (physical or chemical), knowledge, opinions and fads, or behavior.
  In any of these cases, of course, the key issue is to understand how the network architecture bears on the reach of the process. This has been studied in different disciplines, ranging from sociology and economics to molecular biology and neurology. By way of illustration of the breadth of setups considered, see e.g. Granovetter (1973) [136], Bikhchandani *et al.* (1998) [31], Chwe (2000) [63], Kauffman (1993) [170], Sporns (2002) [263], and Amaral *et al.* (2004) [9].

Finally, a third element that makes the field of complex networks distinctively interdisciplinary is of a methodological sort. Complex networks do not lend themselves easily to the tools of analysis that have been traditionally used in the social and biological sciences. Their intricate detail makes it virtually hopeless to attempt a microscopic description of their structure. Besides, even if such minute descriptions were at all feasible, it could hardly be expected that most of the interesting properties of the network would depend on them. For example, the robustness of any large-scale communication network, if it is truly complex, cannot hinge upon very small-scale details on its pattern of connections. Or, in a complex (and thus large) social context, the access to relevant information by the agents (say, on job openings or retail prices) should generally depend on "global" features of the social network. Of course, the same would apply if the problem pertained instead to, say, technological diffusion and the objective were to predict the extent to which the population will eventually come to adopt some innovation.

But, what are the global features of the network that should play a significant role in the analysis? The modern theory of complex networks suggests focusing on those network properties that can be suitably described statistically (e.g. average distances between nodes, their distribution of connectivity, etc.). Correspondingly, the tools used in the analysis must also be of a statistical nature and thus only applicable to very large systems. Statistical physics has a long and fruitful tradition in this type of analysis. It should therefore come to no surprise that much of the methodology used in the analysis of complex networks bears the trace, if not the evident imprint, of this origin. This, of course, reinforces the interdisciplinary character of the field. But it also has an unfortunate side effect: it raises the "entry cost" for outsiders, rendering it more difficult that research becomes a truly interdisciplinary endeavor.

Such a methodological convergence notwithstanding, the different areas of application in the field of complex networks naturally maintain many specificities of their own. This applies, in particular, to the economic and other social networks that will concern us here, for which the distinctive characteristics of socioeconomic environments cannot be ignored. Most importantly, it must be recognized that individual agents (the "nodes" of the social network) have their behavior shaped by what are their prevailing objectives and expectations. The latter, in turn, are sharply affected by the social network, which implies that agents' behavior must be conceived and modeled as embedded in the social network. This leads us to the so-called *issue of network embeddedness*, whose importance was forcefully stressed in a very influential paper by the sociologist Mark Granovetter (1985) [139]. We elaborate on this issue and illustrate its implications in the following subsection.

1.2 SOCIOECONOMIC NETWORKS AND THE ISSUE OF EMBEDDEDNESS

In his appraisal of mainstream economic analysis, Granovetter (1985) [139] sharply criticized it for having traditionally circumscribed to two polar paradigms: *markets* and *hierarchies*. He argued that markets, on one hand, are modeled as an extremely under-socialized setup where social relations play virtually no role. Concerning hierarchies, on the other hand, he suggested that there is an analogous de-socialization, but now resulting, ironically, from an