THE SCIENCE OF COMPLEX NETWORKS

Guido Caldarelli
Centro SMC CNR-INFM, Dipartimento di Fisica, University of Rome "Sapienza", Piazzale Aldo Moro 5 00185 Rome, Italy

Keywords: Complex Networks, small-world effect, Scale-free degree distribution

Contents

Summary

In this theme-level chapter we introduce the subject of complex networks and we present the structure of the associated topic level chapters that range from social science to biology and finance. We start by considering the mathematical foundations of networks and we then move to an overview of the various applications, some of which will be described in detail in the topic-level contribution attached. A glossary of terms

used and appendices on the mathematical notions are also present.

1. Introduction

The field of complex network exploded since the 1990s, the number of publications in a variety of different areas has grown exponentially and practically, and every discipline started to recognize the presence of these mathematical structures in its area of research. Actually almost any system from the nowadays traditional example of the Internet to complex patterns of metabolic reactions can be analyzed through the graph theory. In its simplest and non rigorous definition a graph is a mathematical object consisting of a set of elements (vertices) and a series of links between these vertices (edges). This is of course a very general description, and as any mathematical abstraction, the idea is to discard many of the particular properties of the phenomenon studied. Nevertheless, this modeling is remarkably accurate for a variety of situations. Vertices can be persons related by friendship or acquaintances relations. Vertices can be proteins connected with one another if they interact in the cell. Networks have always existed in Nature of course, but it is fair to say that given the present technological explosion, they became more and more important. Starting from the Internet the web of connections between computers we started to link and share our documents through web applications and we start to get connected with a number of persons larger than usual. It is this revolution in our daily habit that made natural thinking of networks in science and research. Once this has been realized it became natural to see the cell as a network of molecular events from chemical reactions to gene expressions. The point is to establish if this new perspective can help researchers in finding new results and by understanding the development of these phenomena and possibly control their evolution. We believe that this is the case and in the following we shall provide the evidence of that. Together with applications there are of course true scientific questions attached to network theory. Consider the various ways in which the edges are distributed among the vertices: even by keeping the number of edges and vertices constant we have many different patterns possible. Interestingly some features used to describe these shapes are not related to the particular example considered, but instead they are universal. That is to say they can be found in almost any network around.

Among the most important examples for their practical applications are:

1. The robustness of the graph (how long it takes before splitting in two or more parts when edges are sequentially removed),
2. The average distance between two points chosen at random,
3. The number of edges per vertex (degree).

One important application of network theory is that it can provide a new language for the analysis of complex structures. By using this mathematical representation it becomes now possible to have a quantitative description of some characteristics of social systems. It is very easy to model people through vertices and use edges to describe the links between them. An edge can be any interaction between two people, for example to know each other, to work together or to be part of the same group (school, gym etc.). Despite the simplicity of this representation, it works remarkably well when we want to study global properties as the evolution of groups, or the evolution of a virus or opinion formation within a community. You can consider even more complicated situations: friendship or love, in this case the edges are not symmetrical, since the link can be present only one way. This can be described easily considering a directed graph, where the links (as well as one-way streets) may be paths in one direction only.

From these considerations it is clear how complex networks can be considered as the right framework to visualize and operate on complex systems. While several definitions are available from them we can now define them as systems consisting in a large number of parts whose interaction produce some global properties totally unexpected. A similar situation happens in the study of condensed matter: from an atom of an element it is impossible to predict the feature of the macroscopic sample. For example it is impossible to predict if the material will be an insulator or a conductor. These ideas have been introduced by Nobel Laureate PW Anderson [1972] in a seminal paper that became a classic of the science of complexity. Complex networks fit exactly into this definition since the situation is very similar. When combining several servers and connecting them to each other one obtains a structure like the Internet that is characterized by long-range correlations and inherently self-similar. As shown in Figure 1, even a small number of nodes is enough to distinguish the behavior of a hub with respect to the behavior of a peripheral node. A structure similar to that can be spotted essentially unchanged in a large number of systems in a variety of fields, among which finance, biology, and computer science not to mention the various social systems studied so far. In all these cases, beyond the particular conditions, we find the same statistical properties in the topological quantities describing the graph. In the following we will see in detail the experimental evidence of these features and we shall discuss the meaning of these mathematical quantities together with the most recent theoretical understanding of these phenomena. For a more extensive description of the mathematics of complex networks and experimental evidence, see Caldarelli [2007].

2. The Properties of Complex Networks

There are a number of properties that distinguish a complex network from a random graph. The reasons that caused these different properties in real cases are currently still being studied, although the description is clear enough. In principle to fully characterize a graph just give the list of summits with all the different connections. But this information is not always possible (the Web is made up of billions of pages and even more links that vary while you measure), or useful (even a thermodynamic system such as a gas does not describe giving the positions and speeds of each particle, rather use macroscopic quantities as the volume it occupies or the temperature at which it is located). Just as in thermodynamics, you choose to use the global properties given by statistical considerations. This allows you to group together graphs in different classes with properties similar to statistics. These properties not only characterize "statically" one network over another, but also allow you to determine the response of the network to external stresses, giving a description that is "dynamic"

2.1. The Small-World Effect

You can begin to take the first steps in this direction from systems where we can clearly

define the nature of the link. The first experiment of this kind was carried out by American sociologist S. Milgram [1967] that asked people drawn at random from the telephone Obama (Nebraska) to get there a message to a financial trader in Boston (Massachusetts). The rule of the era was that the messages were only to people you know on a personal basis. This means that people should know the recipient or other persons who were in some way more "close" to this person in Boston. Despite the distance and the difference in social environment, on average only a few steps were sufficient to reach the destination of your choice. This result is then passed into popular culture through the comedy (later became a film with W. Smith) "Six degrees of separation." The same concept is also the basis of a very popular game in American colleges. A player says the title of a film, following which the player has to find another film that has at least one player in common with the first. A player wins if he/she finds a movie where Kevin Bacon played. The nature of the link is clear, you can move from one film to another only if they have at least one player in town, and if the players are good enough, regardless of the film starting in a few steps you arrive at your destination. In an entirely different situation is quantified by a similar magnitude as the number of known Erdős. Paul Erdős, a Hungarian mathematician among the most prolific of our time, published about 1,500 scientific articles. All of its 511 co-authors are at 1 distance from him. The co-authors of these co-authors (as long as they are not already in the first list) are a distance from 2 from Erdős and so on.
This property, which often ends up in the systems that we have the properties of "small-world" (the world is small) and is one of the properties that characterize complex networks.

-
-
-

Bibliography

Anderson P.W. (1972) “More is Different” *Science*, **177**, 393-396. [The seminal paper on Complexity Science]

Barabási A.-L., Albert R. (1999) “Emergence of Scaling in Random Networks” *Science* **286,** 509-512 [The first paper that presents the famous BA model]

A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. (2004) “The architecture of complex weighted networks” Proc. Natl. Acad. Sci. USA, **101** 3747.[A paper that extends the property of complex networks to the case where we deal with weights on vertices/edges]

Brockmann D., Hufnagel L., Geisel T. (2006) “The scaling laws of human travel” *Nature* **439,** 462-465 [This paper focuses on the statistical properties of the length and duration of travel in USA]

Buchanan M. (2003) “Small World: Uncovering Nature’s Hidden Network” Weidenfeld and Nicolson, London [A Dissemination book written by a journalist on the network in our society].

Caldarelli G. , Capocci A., De Los Rios P., Muñoz M-A. (2002) “Scale-free Networks from varying vertex intrinsic fitness” *Physical Review Letters* **89**, 258702 [A paper on an alternative model of graph creation, based on the differences between the various vertices].

Colizza V., Barrat A., Barthelemy M., Vespignani A. (2006) "The role of the airline transportation network in the prediction and predictability of global epidemics" *Proceedings of the National Academy of Sciences (USA)* **103** 2015-2020 [Clarifies the small world properties of the world system of flight. From this property the authors study the epidemics diffusion of viruses].

Erdős P., Rényi A. (1959) "On Random Graphs" *Publicationes Mathematicae Debrecen*, **6** 290-297 [The first paper where the model of Random Graph is introduced].

Ginsberg J., Mohebbi M.H., Patel R.S., Brammer L., Smolinski M.S., Brilliant L. (2009) "Detecting influenza epidemics using search engine query data" *Nature* **457**, 1012-1014 [A study of the propagation of flu, from the number of queries related to "flu" submitted to Google].

González M.C., Hidalgo C.A., Barabási A.-L. (2008) "Understanding individual mobility patterns" *Nature* **453,** 479-482 [A study on the habit of individuals in USA as reported from mobile telephone data].

Guimerà R., Mossa S., Turtschi A., Amaral L.A.N. (2005) "The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles" *Proceedings of the National Academy of Sciences* (USA) **102** 7794–7799 [As above a study on the flight transportation network that allow to detect particular areas where disease can become epidemic].

Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E. & Aberg, Y. (2001) "The web of human sexual contacts." *Nature* **411**, 907–908 [Experimental evidence of the scale-free nature of the number of sexual intercourses. This analysis is crucial in order to define a good strategy to cure sex-transmitted diseases propagation]

Milgram, S. (1967) "The small world problem" *Psychology Today* **1**, 60-67 [The seminal paper where the small-world effect is introduced].

Newman M.E.J., Girvan M.,(2003) "Finding and evaluating community structures in networks" *Physical Review E* **69**, 026113 [One of the best paper on the study and determination of communities in networks].

Pastor-Satorras, R. & Vespignani, A., (2001) "Epidemia Spreading in Scale Free Networks" *Physical Review Letters* **86**, 3200–3203 [The first paper where it is understood that epidemics on a scale-free networks can propagate without a minimum number of persons infected]

Biographical Sketch

Guido Caldarelli is currently Associate Professor in the Institute of Complex Systems in the Department of Physics of the University of Rome "Sapienza" Italy. The institute is part National Research Council (CNR) of Italy.

He got his degree in physics in the Department of Physics of the same University in 1992 working with L. Pietronero and A. Vespignani. He then moved to SISSA/ISAS in Trieste where he got the PhD in Statistical Physics in 1996 working on Self-Organized Criticality with A. Maritan. He has been postdoc in the Department of Physics in the University of Manchester with A. McKane and in TCM Group in the University of Cambridge with R. Ball. During his scientific activity in Rome he has also been visiting professor in the École Normale Supérieure in Paris, and in the Department of Physics of the University of Barcelona.

After the studies on fractal growth and self-organized criticality he moved his research on the analysis of scale-free networks. On this topic he published a textbook and he coordinated a European Project (http://www.cosinproject.org).