<u>Contribution :</u> 2005 NAE Frontiers of Engineering

Complex Networks:
Ubiquity, Importance and Implications

Alessandro Vespignani

School of Informatics and Department of Physics, Indiana University, USA

In recent times, the increased power of computers and the informatics revolution have made possible the systematic gathering and handling of data sets on several large scale networks, allowing the detailed analysis of their structural and functional properties. In particular, mapping projects of the World-Wide Web and the physical Internet offered the first chance to study the topology and traffic of large scale networks. Gradually other studies followed describing networks of practical interest in social science, infrastructures analysis and epidemiology. The study of these systems involves researchers from many different disciplines and has led to a shift of paradigm in which the *complexity* of networks has become the central issue in their characterization, modeling and understanding[1].

- Ubiquity

In very general terms a network system is described as a graph whose nodes (vertices) identify the elements of the systems and the set of connecting links (edges) represent the presence of a relation or interaction among these elements. With such a high level of generality it is easy to perceive that a wide array of systems can be approached within the framework of network theory. In the first instance we can provide a rudimentary taxonomy of the real world networks. Two main different classes can be related with infrastructure systems and natural or living systems, respectively. Each one of these classes can be further divided in different subgroups. Networks belonging to the class of natural systems find a differentiation in the subgroups of biological networks, social system networks, food-webs and ecosystems, just to mention a few. For instance, biological networks refer to the complicate set of interactions among genes, proteins and molecular processes that regulate the biological life, while social networks concern relations between individuals such as family relationships, friendship, business and many others[2, 3].

In turning our attention to infrastructure networks we can readily individuate two main subcategories. The first one contains *virtual* or *cyber* networks. Those networks exist and operate in the digital world of cyberspace. The second one includes physical systems such as energy, transportation or healthcare networks. This is of course a rough classification since there are many interrelations and interdependencies existing among physical infrastructure networks as well as between physical and digital networks. The Internet, for instance, is a kind of hybrid network in which the cyber features are mixed with the physical features. It is comprised of physical objects such as routers -the main computers which allow us to communicate- and transmission lines, the cables which connect the various computers. On top of this physical layer, we have a *virtual* world made of software that may define different networks such as the world-wide-web (WWW), the e-mail network and Peer-to Peer networks. These networks are the information transfer media for hundreds of millions of users and, similarly to the physical Internet, have grown to become enormous and intricate networks as the result of a self-organized growing process. Their dynamics is the outcome of the interactions among the many individuals forming the various communities and in this sense they are a mixture of complex socio-technical aspects. Further examples of this kind can be found in the world-wide airport and power distribution networks where physical and technological constraints cooperate with social, demographic and economical factors.

- Complexity

The questions "where do we find complex networks" and "why we define them *complex*" imply the distinction of what is "*complex*" and what is the merely complicated. This distinction is a critical one because the characteristic features and the behavior of complex systems differ significantly from those of merely complicated systems. A minimal definition of complexity may involve two main features: i) the system exhibits complications and heterogeneities that extend virtually on all scales allowed by the physical size of the system; ii) these features are the spontaneous outcome of the interactions among the many constituent units of the system, i.e. we are in the presence of an emergent phenomenon.

It is easy to realize that the WWW, the Internet, the airport network are all systems which grow in time by following complicate dynamical rules and without a global supervision or blueprint. The same can be said for many social and biological networks. All these networks are self-organizing systems, which at the end of the evolution show an emergent architecture with unexpected properties and regularities. However, if complexity resides in the emergence of complications at all scales, one might then wonder where we find a signature of complexity in real networks. A first clue is provided by the high level of heterogeneity in the degree of vertices, i.e. the number of connection k of each vertex. This feature is easily depicted by the visual inspection of the airport networks and the "hub" policy that almost all airlines adopt. The same arrangement can be easily perceived in many other networks where the presence of "hubs" is a natural consequence of different factors such as popularity, strategies and optimization. For instance, in the WWW some pages become so popular to be pointed by thousands of other pages, while in general a majority of documents is almost unknown. The presence of hubs and connectivity ordering turn out to have a more dramatic manifestation than initially thought[4], yielding a degree distribution $P(k)$ with heavy-tails often approximated by power-law forms $P(k)\sim k^{-g}$. The peculiar fact about a distribution with a heavy tail is that the average behavior of the system is not typical. In more mathematical terms the heavy-tail property translates in a very large level of degree fluctuations with the divergence of the standard deviation of the degree distribution, limited only by the finite size of the systems[4,5]. We are thus in the presence of structures whose fluctuations and complications extends over all possible scales allowed by the physical size of the systems, i.e. we are facing complex systems.

-Importance

Heavy tails and heterogeneity appear to be common to a large number of real world networks, along with other complex topological features such the presence of communities, motifs, hierarchies and modular ordering. The evidence that a complex topology is shared by many complex evolving networks cannot be considered as incidental. Rather, it points to the possibility of some general principle that can possibly explain the emergence of this architecture in such different contexts. In this perspective, it becomes particularly relevant to have a theoretical understanding that might uncover the very general principles underlying the networks formation. When looking at networks from the point of view of complex systems, the focus is placed on the microscopic processes that rule the appearance and disappearance of vertices and links. The attempt to model and understand the origin of the observed topological properties of real networks therefore results in a radical change of perspective that shifts the focus on predicting the large scale properties and behavior of the system on the basis of the dynamical interactions of its many constituent units. For this reason, in the last years a very intense activity has been focused on the development of dynamical models for networks, eventually generating a vast field of research whose results and advances provides new techniques to approach conceptual and practical problems in the field of networked systems[1,4].

- *Implications*

The advances obtained in the understanding of large complex networks have also generated an increased attention towards the potential implication of complex properties with respect to the most important questions concerning their engineering, optimization and protection. These problems are emerging as fundamental issues whose relevance of goes beyond the usual basic science perspective. For instance, the complexity of networks has important consequences in the empirical analysis of the robustness in front of failures and attacks.

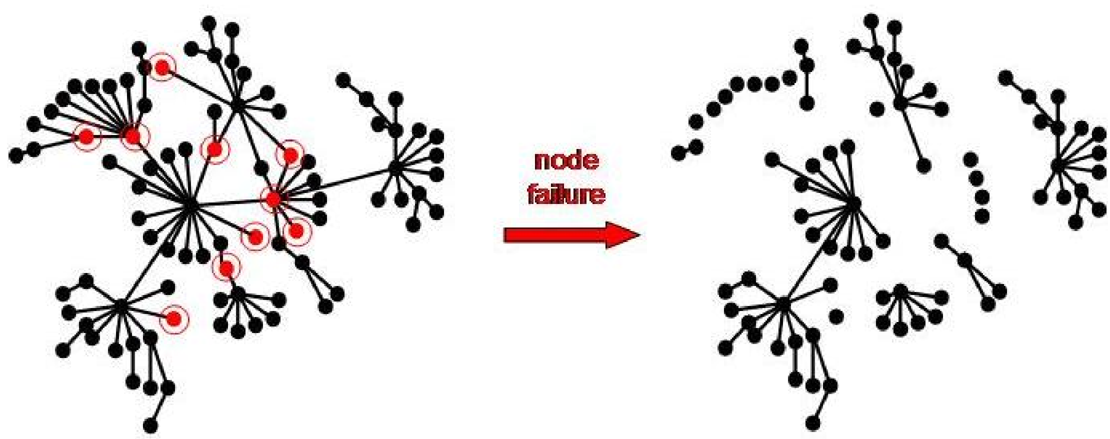

Figure 4: Analysis of the topological resilience to attacks and failures. Encircled nodes in the network (left) are removed as a result of malfunctioning or attack. After damage (right) the network consists of several fragmented components.

A natural question to ask in this context concerns the maximum amount of damage that the network can take, i.e. the *threshold* value of the removal density of vertices above which the network can be considered compromised. Contrary to regular networks, heavy-tailed networks present two faces in front of component failures: they are extremely robust to the loss of a large number of randomly selected vertices, but extremely fragile in response to a targeted attack[6].

Another basic example of the impact of complex network studies is provided by the insights obtained on the properties of disease spreading in highly heterogeneous networks[7]. Indeed, the presence of heavy-tailed connectivity patterns changes dramatically the epidemic framework usually obtained in more regular networks. In the latter case it is possible to show on a general basis that if the epidemic spreading rate – roughly speaking the disease transmission probability- is not larger than a given threshold value, the epidemic dies in a very short time. On the contrary in scale-free networks, whatever the spreading rate, it exists a finite probability that the infection will pervade the system with a major outbreak or a long lasting steady state, i.e. heavy-tailed networks lack of any epidemic threshold[8]. Interestingly, the absence of any epidemic threshold corresponds also to a general inadequacy of uniform immunization policies. On the other hand, it is possible to take advantage of the connectivity pattern of heavy-tailed networks by showing in mathematical terms that the protection of just a tiny fraction of the most connected individuals raises dramatically the tolerance to the disease of the whole population[9].

Finally, complexity features affect also in the dynamics of information or traffic flow taking place on their structure. The resilience and robustness of networks is a dynamical process, which should take into account the time response of elements to different damage configurations. For instance, after any router or connection fails, the Internet responds very quickly by updating the routing tables of the routers in the neighborhood of the failure point. While in general this adaptive response is able to circumscribe the damage, in some cases failures may avalanche through the network, causing far more disruption that one would expect from the initial cause[10]. This is typical of complex systems where emergent properties imply that events and information spread over a wide range of length and time scales. In other words, small perturbations have a finite probability to trigger a system-wide response, the so-called critical behavior. This happens through chains of events that eventually may involve a large macroscopic part of the system and, in some cases, lead to a global failure. It is important to realize that in large networked system this property is inherent to the system's complexity and it is not changing by using local reinforcements or technological updates. We can vary the proportion of small or large events, but we have to live with appreciable probabilities for very large events: we must deal with the inherent complexity of the real world.

- *References*

[1] A.-L. Barabási "Linked": Perseus Books Group 2002; S.N. Dorogovstev and J. F. F. Mendes, "Evolution of Networks", Oxford University Press: Oxford, 2003; R. Pastor-Satorras, and A. Vespignani "Evolution and structure of the Internet", Cambridge University Press: Cambridge, 2004; M.E.J. Newman,"Structure and function of complex networks", *SIAM Review* 45, 167-256 (2003).

[2] M., Castells "The Rise of the Network Society", Volumes 1 and 2, Oxford: Blackwell, 2000.

[3] S. Wasserman and K. Faust, " Social Network Analysis : Methods and Applications", Cambridge University Press: Cambridge, 1994.

[4] A.-L. Barabasi and R. Albert "Emergence of scaling in random networks", *Science* **286**, 509-511 (1999); R. Albert and A.-L. Barabasi, "Statistical mechanics of complex networks", *Rev. Mod. Phys.* **74**, 47-97 (2002).

[5] L.A.N. Amaral et al ., "Classes of small world networks" , Proc. Natl. Acad. Sci (USA) 97, 11149-11152 (2000).

[6] R. Albert et al., "Error and attack tolerance of complex networks", *Nature* **406**, 378-382 (2000); R.Cohen et al. "Resilience of the Internet to random breakdowns" Physical Review Letters **85**, 4626-4629 (2000).

[7] F. Lijeros et al., "The web of human sexual contacts", *Nature* **411**, 907-908 (2001); A. Schneeberger at al., "Scale-free networks and sexually transmitted diseases", *STD* **31**, 380-387 (2004).

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

[9] Details in R. Pastor-Satorras and A. Vespignani, "immunization of complex networks" *Physical Review* E **65**, 036104 (2001); Z. Dezso and A.-L. Barabasi,"Halting viruses in scale-free networks" *Physical Revew* E **65**, 055103 (2001); R. Cohen et al., "Efficient immunization strategies for computer networks and populations*", Physical Review Letters* ***91****, 247201 (2003).*

[10] A.E. Motter and Y.C.Lai, "Cascade based attacks on complex networks", Phys.Rev. E **66**, 065102 (2002); Y. Moreno et al. "Critical load and congestion instabilities in scale-free networks", *Europhys. Lett.* **62**, 292 (2003); E.J.Lee et al. "Robustness of the avalanche dynamics in data packet transport on scale-free networks", Phys. Rev. E **71**, 056108 (2005).