IJCSS International Journal of Complex Systems in Science

Int. J. Complex Systems in Science
vol. 1 (2011), pp. 78–82

Uncertainty in Complex Networks

Massimiliano Zanin[1,†]

[1] *INNAXIS Foundation & Research Institute, Madrid, Spain*

Abstract. Complex networks have been a valuable tool to analyze complex systems, and to understand their structure and inner dynamics. They are usually considered as deterministic; the existence of a link is an initial information, or is defined by the evolution of some dynamics: but this information is always known at the time of analyze the graph. In this work, we introduce a new approach to define complex networks which includes uncertainty inside the adjacency matrix.

Keywords: Complex networks, uncertainty
MSC 2000: 05C82

1. Introduction

Complex Networks [1] have been the focus of many researches in the last years, which have covered both theoretical aspects and applications for different real systems [2]. The usual way to describe the topology of a network is through an adjacency matrix A of $n \times n$ elements, where n is the number of nodes. Each entry $a_{i,j}$ $(i, j = 1 \dots n)$ is equal to 1 when a direct connection exists between nodes i and j, and zero otherwise. To this adjacency matrix, a matrix of weights W is usually added, where elements $w_{i,j}$ are real numbers attached to the corresponding links and representing some kind of weight, length or capacity. In spite of the flexibility of this mathematical tool to handle different kinds of structures, there are real-world systems whose characteristics cannot be easily described by complex networks. For example, consider the transportation network of a city. At a first approach, it is possible to create a complex network from streets and calculate geodesic paths and other measures from it [3]. Nevertheless, it is not possible to ensure that all streets will be practicable at a given time; in other words, in a *strategic level* the only information available is about the probability to have a street closed because of accidents or other random events.

This contribution tackles uncertainty inside complex networks, by developing the mathematical foundation for a probabilistic adjacency matrix, and presenting some results related to standard network models.

2. Probabilistic Complex Network

The concept of probabilistic networks was already introduced by Ahnert *et al.* [4]. In their work, some measures commonly used to characterize complex networks topology have been transposed within this probabilistic approach; nevertheless in this work we focus on the relation between length of a path and the probability of existence of the same, for its implication in the study of transportation networks. Elements of the new probabilistic adjacency matrix $\mathcal{A}$ are defined in the range $0 \leq a_{i,j} \leq 1$: that is from zero (when that connection never activates) to one (the link is always present). Values between both extremes represent the probability for a link to exist in the network.

Paths in complex networks are sequences of links connecting two nodes; except in the trivial situation where all elements of $\mathcal{A}$ are zero or one, in this framework we cannot tell nothing but the probability for any path to exist. Such probability is given by the probability of existence of each link included in that path:

$$p_{path} = \prod_{l \in path} a_l \tag{1}$$

The concept of *geodesic path*, that is the shortest path connecting two nodes, must now be complemented with the *most probable path*, or the path with maximal probability of existence. In most cases, both paths are expected to be different: for example, it is usual to find a shortest path crossing the congested center of a city, and a longer path with less uncertainty. Finding the path between two nodes with the higher existence probability is an easy task, and it can be accomplished with any standard path-searching algorithms, e.g. Dijkstra's. First, it must be noted that:

$$\sum_i \ln x_i = \ln \prod_i x_i \tag{2}$$

If the elements of a new *Logarithmic Adjacency Matrix* $\mathcal{L}$ are defined as

$$l_{i,j} = -\ln a_{i,j} \tag{3}$$

the problem of maximizing the probability of a path (i.e. maximizing the product of the probabilities in $\mathcal{A}$) can be transformed into a standard problem of minimizing the cost of the path (i.e. minimizing a sum), where the cost of each step is given by the elements of $\mathcal{L}$.

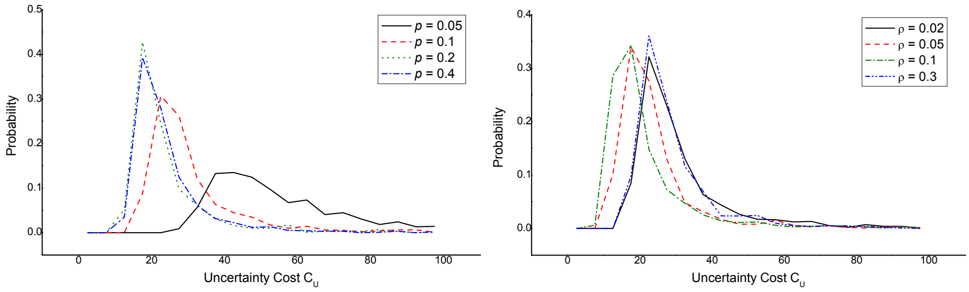

Figure 1: Histograms of the Uncertainty Cost C_u in random networks of 100 nodes, as a function of (Left) the probability p for a node to be attached to another node ($\rho = 0.2$), and (Right) the standard deviation of link probabilities ($p = 0.1$).

2.. 1 The Cost of Uncertainty

The existence of a short and insecure path, and of a secure but longer path, results in the problem of estimating the *cost* generated by uncertainty. In some situations, exact information can be obtained about this cost; for instance, in the network of streets of a city, one can compare the cost of being trapped in a traffic jam, with the cost of taking a longer route (fuel burnt).

Nevertheless, this approach is not always feasible. If only the topology of the system is known, the cost can be estimated as the increment of length necessary to raise the existence probability in an arbitrary quantity. That is, this cost of uncertainty can be defined as the cost we need to pay in order to reduce such uncertainty. Mathematically, if d is the length of a path, p its existence probability, and S and M are the shortest and the most probable paths respectively, the cost is given by:

$$C_u = \frac{d_M - d_S}{p_M - p_S} \tag{4}$$

Some results for C_u in different network topologies are presented. In Fig. 1 are shown the histograms for random networks of 100 nodes. In Fig. 1 Left are represented the results for different number of links: p is the probability to connect each couple of nodes, so that the result is a Erdös and Rényi (ER) random graph $G^{ER}_{100,p}$ [5]; the existence probability of each link of the network is taken from a normal distribution centered in 0.5 and with standard deviation 0.1. Fig. 1 Right is constructed by fixing $p = 0.1$ and by varying the standard deviation. Results clearly show that a reduction in the number of connections inside the network leads to an higher mean Uncertainty Cost.

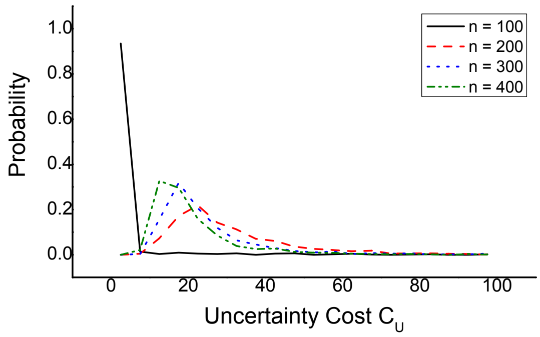

Figure 2: Histogram of the Uncertainty Cost C_u for Scale-Free networks for different numbers of nodes.

Another important type of network, which has been encountered in several real systems, are Scale-Free networks, whose degree distribution is shaped according to a power law in the form $P(k) \approx Ak^{-\gamma}$. In Fig. 2 are shown the histograms of the Uncertainty Cost for several Scale-Free networks created with the Barabási-Albert growing mechanism [6]. Each one of them is the result of 100 growth iterations: therefore, for instance, in the network with $n = 200$, two nodes were added at each iteration. It is interesting to note how the probability of get an Uncertainty Cost of zero, for $n = 100$, is very relevant. It must be observed that, in this case, just one node is added at each iteration: therefore there is usually only one path connecting two nodes, which is therefore the shortest and the most secure one. It is also interesting to note that distributions for Scale-Free networks show a long-tail effect, which decreases when more nodes are added at the same time. These results could have important consequences in real systems; it is known that Scale-Free networks are more vulnerable to targeted attacks: but they have also a smaller resilience against uncertainty.

3. Conclusions

In this contribution, a new definition of the adjacency matrix of a network is introduced, with the aim to include uncertainty in its topology. With this approach, it is possible to estimate the *cost of uncertainty* of a network, as the cost (in term of path length) that should be payed to obtain a more secure path, when traveling between two of its nodes. This concept should allow a better characterization of real systems, specially of those where uncertainty is an important aspect of their dynamics: for instance, any transportation network, like the air transport system.

References

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Physics Reports* **424**, (2006). M. E. J. Newman, *SIAM Review* **45**, 67–256 (2003).

[2] L. da F. Costa, O. N. Oliveira Jr., G. Travieso, F. A. Rodrigues, P. R. V. Boas, L. Antiqueira, M. P. Viana, and L. E. Correa da Rocha, *arXiv:0711.3199v2 [physics.soc-ph]*, (2008).

[3] M. Rosvall, A. Trusina, P. Minnhagen, and K. Sneppen, *Physical Review Letters* **94**, 028701 (2005). P. Crucitti, V. Latora, and S. Porta, *Physical Review E* **73**, 036125 (2006).

[4] S. E. Ahnert, D. Garlaschelli, T. M. A. Fink, and G. Caldarelli, *Physical Review E* **76**, 016101 (2007).

[5] P. Erdös, and A. Rényi, *Publicationes Mathematicae* **6**, (1959).

[6] R. Albert, and A.-L. Barabási, *Reviews of Modern Physics* **74**, (2002).