

## ANALYZING COMPLEX NETWORKS

ANGEL GARRIDO

ABSTRACT. Our paper analyzes some new lines to advance on quickly evolving concepts, as the so-called Symmetry and Entropy, on graphs and networks in general. So, it will be very necessary to analyze the mutual relationship between some fuzzy measures, with their interesting applications; in particular, when we works on Complex Networks and Systems.

2000 *Mathematics Subject Classification*: 03B52, 03E72, 05C72, 23E10.

*Keywords*: Fuzzy Logic; Fuzzy Set Theory; Mathematical Analysis; Fuzzy Measures; Fuzzy Graph Theory; Complex Networks.

## 1. INTRODUCTION

We need to analyze here some very interrelated concepts about a graph [7, 15], as may be the case of Symmetry / Asymmetry degrees [14], their Entropies, Clustering Coefficient, etc. It may be applied when we study the different types of Systems, and in particular, on Complex Networks [3, 5, 16, 17].

A *system* can be easily defined as a set of components functioning together as a whole [4]. A systemic point of view allows us to isolate a part of the world, and so, we can focus on those aspect that interact more closely than others.

The *Network Science* is a new scientific field that analyze the interconnection among diverse networks, as for instance, on Physics, Semantics, and so on [1-5]. Among its developers, we may remember Duncan Watts, with the *Small-World Network*, or Albert-László Barabasi and Réka Albert, who developed the *Scale-Free Network*. In his work, both authors have found that the websites that constitutes the network of the WWW have very interesting

mathematical properties. So, Network Theory is a very quickly expanding area of Computer Sciences, and may be considered as a part of Graph Theory.

Complex Networks are everywhere. Many different phenomena in Nature can be modeled as a network [3], such as

- *Brain structures*. The brain as a network of neurons (their nodes), connected by synapses (their edges).

- *Social interactions*, or the World Wide Web (WWW).

All such systems can be represented in terms of nodes and edges.

- On the Internet, the nodes represent routers and the edges are represented by wires, or physical connections between them.

- In transport networks, the nodes can represent the cities and the edges the roads that connect them. These edges can have weights. Such networks are not random.

## 2. GRAPH THEORY

The topology of very different networks are very close [7, 15]. They are rooted in the Power Law, with a scale free structure. How can very different complex systems have the same underlying topological features? Searching the hidden laws of these networks [11, 12, 16-18], modeling and characterizing them are the current lines of research. It has emerged as a primary tool for detecting numerous hidden structures in different information networks, including Internet graphs, social and biological networks, or any graph representing relations in massive data sets. The analysis of these structures is very useful for introducing concepts such as Graph Entropy and Graph Symmetry. We consider a functional on a graph,  $G = (V, E)$ , with  $P$  a probability distribution on its node set,  $V$ . The mathematical construct called as *Graph Entropy* will be denoted by  $GE$ . Such a function is convex. It tends to plus infinite on the boundary of the non-negative orthant on the Euclidean  $n$ -space. And monotonically to minus infinite along rays departing from the origin. So, such minimum is always achieved and it will be finite.

The entropy of a system ( denoted by  $H$  ) represents the amount of uncertainty one observer has about the state of the system [6, 14].

The simplest example of a system may be a random variable and will be shown by a node into the graph, which then gives us the representation of the mutual relationship between them by the corresponding edges.

*Information* measures the amount of correlation between two systems, and it will be reduced to a difference in entropies [5, 14, 16]. So, the Entropy of

a Graph is a measure of graph structure, or lack of it. Therefore, it may be interpreted as the amount of information, or the degree of surprise communicated by a message. And, since the basic unit of Information is the bit, the Entropy may also be viewed as the number of bits of "randomness" in the graph validating that *the higher the entropy, the more random the graph is*.

*Symmetry*, in a system, means invariance of its elements under conditions (group) of transformations. When we take Network Structures, it means invariance of adjacency of nodes under permutations of the node set [7, 14]. The graph isomorphism is an equivalence, or equality, as relation on the set of graphs. Therefore, it partitions the class of all the graphs into equivalence classes.

The underlying idea of *isomorphism* is that some objects have the same structure, if we omit the individual character of their components. A set of graphs isomorphic to each other is denominated as an isomorphism class of graphs. The *automorphism* of a graph,  $G = (V, E)$ , will be an isomorphism from  $G$  onto itself.

The family of all automorphisms of a Graph,  $G$ , is a permutation group on  $V(G)$ . The inner operation of such a group will be the composition of permutations. Its name is very well-known, the *Automorphism Group of  $G$* , and abridgedly, it will be denoted by  $\text{Aut}(G)$ . Conversely, all the groups may be represented as the automorphism group of a connected graph.

The automorphism group is an algebraic invariant of a graph [7, 15]. So, we can say that the automorphism of a graph is a form of symmetry in which the graph is mapped onto itself while preserving the edge-node connectivity. Such an automorphic tool may be applied both on Directed Graphs (DGs), and on Undirected Graphs (UGs).

We will say either *graph invariant*, or *graph property*, when it depends only on the abstract structure, not on graph representations, such as particular labelings, or drawings of the graph. So, we may define a graph property as every property that is preserved under all the possible isomorphisms of the graph.

Therefore, it will be a property of the graph itself, not depending on the corresponding representation. The semantic difference also consists in its character: qualitative or quantitative one. When we say "the graph does not possess directed edges", this will be a property, because it result a qualitative statement. Whereas when we say "the number of nodes of degree two in such graph", this would be an invariant, because it is a quantitative statement.

From a strictly mathematical viewpoint, a graph property can be interpreted as a class of graphs.

Being composed by the graphs that have in common the accomplishment of some conditions. Hence a graph property can also be defined as a function whose domain would be the set of graphs, and its range will be the bi-valued set composed by two usual logical options, True and False, according to which a determinate condition is either verified or violated for the graph [7].

A graph property is called *hereditary*, if it is inherited by their induced subgraphs. And it is *additive*, if it is closed under disjoint union. So, for instance, the property of a graph to be planar is both additive and hereditary. Instead of this, the property of being connected is neither [14].

The computation of certain graph invariants may be very useful indeed, with the objective of discriminate when two graphs are isomorphic, or rather non-isomorphic. The support of these criteria will be that for any invariant at all, two graphs with different values cannot be isomorphic between them. But nevertheless, two graphs with the same invariants may or may not be isomorphic between them. So, we get to the notion of *completeness* [16-18].

A digraph contains no cycle if and only if all eigenvalues of its adjacency matrix are equal to zero.

It is provable that every group is the automorphism group of a graph. If the group is finite, the graph may be taken to be finite. And György Pólya observed that not every group is the automorphism group of a tree.

The structural information content will be the entropy of the underlying graph topology [7, 14]. A network is said to be *asymmetric*, if its automorphism group reduces to the identity group. I.e. it only contains the identity permutation. Otherwise, the network is called *symmetric*. I.e. when the automorphism group has elements different from the identity.

### 3.COMPLEX NETWORKS

As a previous concept [4, 5, 11-13], we need to introduce the *Clustering Coefficient*, in both versions, Global and Local. The *Characteristic Path Length* defines the distance from every node to every other node. It will be calculated by the median of the shortest path from every node to every other node. And the *Diameter* defines the maximum possible distance between all the pair of reachable nodes.

The *Global Clustering Coefficient* ( denoted by  $C$ ) is based on triplets of nodes, and can be obtained through the number of closed triplets over the

total number of triplets, both closed and open. It would be also defined as the mean of the clustering indices of all the nodes in the graph. Being computable through the neighbors of the node, by searching the number of links (edges) amongst them. Because the ratio of the number of existing links to the number of possible links gives the clustering index of the node. Whereas the Local Clustering Coefficient will be related to a particular node,  $i$ . It measures the closeness that their neighbors are to being a clique, i.e. a complete graph. The more usual notation is  $C_i$ .

This measure was introduced by Watts and Strogatz, in 1998, to determine whether a graph is a Small-World network. The *degree* of a node  $i$  into a graph may be defined as the number of nodes comprised in its neighborhood. It is the total degree of the corresponding node,  $i$ , i.e. it may be calculated by adding in-degree of  $i$  and out-degree of  $i$ . And finally, the Local Clustering Coefficient for the whole system is the average of the clustering coefficient for each node. So, it is a property of a node into the network, either of the total network. It tells about the connectedness between the neighbors of the nodes. When the neighborhood is fully connected, the value of the clustering coefficient will be equal to one. And it will be a value next to zero, if there are sparse connections in the neighborhood. And with these, all the scale of intermediate cases, [7, 15].

#### 4. STRUCTURAL MODELS

We may distinguish now four of such structural models [9-12]. Thus,

- *Regular Networks*
- *Random Networks*
- *Small - World Networks*
- *Scale - free Networks*

In the Regular Network, each node is connected to all other nodes [6, 7]. I.e. they are fully connected. Because of such a type of structure, they have the lowest path length (L), and the lowest diameter (D), being  $L = D = 1$ . Also, they have the highest clustering coefficient. So, it holds  $C = 1$ .

As related to Random Graphs (RGs), we can say that each pair of nodes is connected with probability  $p$ . They have a low average path length. Moreover, they possess a low clustering coefficient, when the graph is sparse, given that the probability of each pair of neighboring nodes to be connected is precisely equal to  $p$ . The Small-World effect is observed on a network when it has a low average path length. Recall that the now famous "*six degrees of separation*"

is also known as the small-world phenomenon [8]. This initially surprising hypothesis was firstly formulated by the writer Frygias Karinthy in 1929. Then, it was tested by Stanley Milgram in 1967. The subjacent idea [1, 2, 8] is that two arbitrarily selected people may be connected by only six degrees of separation (in average, and it is not much larger than this value). Therefore, the diameter of the graph is not much larger than six. For instance, on social connections.

In the case of the Watts-Strogatz (in acronym, WS) small-world model, it represents a hybrid case between a Random Graph and a Regular Lattice. So, Small-World models share with Random Graphs some common features, such as

- Poisson or Binomial degree distribution, near to the Uniform distribution.
- Network Size: it does not grow.
- Each node has approximately the same number of edges, i.e. it shows a homogeneous nature.

Such models show the low average path length typical of Random Graphs,  $L$  will be equivalent to  $\ln n$ , for  $n$  greater than 1. And also such models gives the usual high clustering coefficient of Regular Lattices, giving  $C$  equivalent to  $.75$ , for  $k$  greater than 1. Therefore, the WS-models have a small-world structure, being well clustered. The Random Graphs coincides on the small-world structure, but they are poorly clustered. This model (Watts-Strogatz) has a peak degree distribution of Poisson type. With reference to the above last model (Scale-Free Network), this appears [2] when the degree distribution follows a Power-Law. In such a case, there exists a small number of highly connected nodes, called , *Hubs*, which are the tail of the distribution.

On the other hand, the greatest majority of the sets of nodes have very few connections, representing the head of such distribution. Such a model was introduced by Albert-Lászlo Barabási and simultaneously by Réka Albert (1999). But there exist certain historical precedents, as George Udny Yule (1925), Herbert Simon (1955) and Derek de Solla Price (1976), among others.

Some of their principal features may be these:

- *Non-homogeneous nature*, in the sense that some (few) nodes have many edges from them, and the remaining nodes only have very few edges, or links.
- As related to the network size, it *continuously grows*.
- And regarding to the connectivity distribution, *it obeys a Power-Law distribution*.

A very notable example of Scale-Free Network may be the World Wide Web

(WWW). It will be described [2, 3, 8, 9, 10] by a collection of sub-networks. In such case, the nodes correspond to their pages, whereas the edges correspond to their hyperlinks. Also related to the Web graph characteristics, we may notice the Scale Invariance, as a very important feature of such very notable complex network.

Community structure (also called *Modularity*) is a very frequent characteristic in many real networks. Therefore, it has become a key problem in the study of networked systems. Giving out its deterministic definition will be a nontrivial problem for the complexity of networks. A concept of modularity  $Q$  can be used as a valid measure for community structure.

Some current models have proposed to capture the basic topological evolution of complex networks by the hypothesis that highly connected nodes increase their connectivity faster than their less connected peers, a phenomenon denoted as *PA* (*preferential attachment*). So, we will find a class of models that view networks as evolving dynamical systems, rather than static graphs [9, 10, 13].

Most evolving network models are based on two essential hypotheses: growth and preferential attachment. The growth suggests that networks continuously expand through the addition of new nodes and links between the nodes. While the preferential attachment states that the rate with which a node with  $k$  links acquires new links is a monotonically increasing function of  $k$ . An important case in community detection is that some nodes may not belong to a single community, and then placing them into more than one group is more reasonable. Such nodes can provide a “fuzzy” categorization, and hence, they may take a special role, such as signal transduction in biological networks [14].

So, I hope to be reached our principal objective, which was simply to provide a very comprehensive perspective on the essential aspects of Graphs and Complex Networks, with some new and interesting results, but ever departing from a Mathematical Analysis point of view.

#### REFERENCES

- [1] R. Albert, and A.- L. Barabási. Statistical Mechanics of Complex Networks, Reviews of Modern Physics, Volume 74, (2002), pp. 47-56.
- [2] A.- L. Barabási, and Bonabeau, Scale-Free Networks, Scient. Am., May 2003, pp. 50-59.
- [3] A.- L. Barabási, Linked: How Everything is Connected to Everything Else, Plume Publ., 2004.

- [4] A. Barrat et al., *Dynamical processes in Complex Networks*, Cambridge University Press, 2008.
- [5] S. Boccaletti et al., *Complex Networks: Structure and Dynamics*, Phys. Rep., Volume 424, (2006), pp. 175-308.
- [6] B. Bollobás, *Random Graphs*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2001.
- [7] S. Bornholdt, and H. G. Schuster, editors, *Handbook of Graphs and Networks: From the Genome to the Internet*. Wiley-VCH, 2003.
- [8] G. Calderelli, *Scale-Free Networks*, Oxford University Press, 2007.
- [9] R. Cohen et al., Resilience of the Internet to random breakdown, Phys. Rev. Letter, Volume 85, (2000), pp. 4626-4242.
- [10] S. Dill et al., Self-similarity in the web, ACM Trans. Intern. Techn., (2002), pp. 205-223.
- [11] S. N. Dorogotsev, and J. F. F. Mendes, *Evolution of Networks*, Adv. Phys. Volume 51, (2002), pp. 1079-1088.
- [12] S. N. Dorogotsev, and J. F. F. Mendes . *Evolution of Networks: From biological networks to the Internet and WWW*. Oxford University Press, 2003.
- [13] S. N. Dorogotsev, A. V. Goltsev, and J. F. F. Mendes, Critical phenomena in Complex Networks, Rev. Mod. Phys., Volume 80, (2008), pp. 1275-1288.
- [14] A. Garrido, Symmetry and Asymmetry Level Measures, Symmetry journal 2010, Volume 2, Issue 2, (2010), MDPI, Basel, Switzerland, pp. 707-721.
- [15] F. Harary, *Graph Theory*, Addison-Wesley, 1975.
- [16] M. Newman, The structure and function of Complex Networks, SIAM Rev., Volume 45, (2003), pp. 167-256.
- [17] M. Newman et al., *The structure and Dynamics of Complex Networks*, Princeton University Press, 2006.
- [18] S. H. Strogatz, Exploring Complex Networks, Nature, Volume 410, (2001), pp. 268-276.

Angel Garrido  
Department of Fundamental Mathematics  
University of UNED  
Senda del Rey 9, 28040-Madrid, Spain  
email: *agarrido@mat.uned.es*