Symmetry of Complex Networks

Angel Garrido

Facultad de Ciencias de la UNED

Abstract

Our paper analyzes some new lines to advance on quickly evolving concepts, the so-called Entropy, or its Symmetry/Asymmetry degrees, on graphs in general. It will be very necessary to analyze the mutual relationship between some fuzzy measures, with their very interesting applications, as may be the case of Graph Entropy and Graph Symmetry; in particular, working on Complex Networks and Systems.

Keywords: Measure Theory, Fuzzy Measures, Symmetry, Entropy.

Mathematics Subject Classification: 68R10, 68R05, 05C78, 78M35.

1. Introduction

We need to analyze here some very interrelated concepts about a graph, such as may be their Symmetry/Asymmetry degrees, their Entropies, etc. It may be applied when we study the different types of Systems; in particular, on Complex Networks [7].

A *system* [8] can be defined as *a set of components functioning together as a whole.* 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.

According Wilson [39], "the greatest challenge today, not just in cell biology and ecology but in all of science, is the accurate and complete description of complex systems. Scientists have broken down many kinds of systems. They think they know most of the elements and forces. The next task is to reassemble them, at least in mathematical models that capture the key properties of the entire ensembles."

Network Science is a new scientific field that analyzes the interconnection among diverse networks, as for instance, on Physics, Engineering, Biology, Semantics, and so on. Between its developers, we may remember to Duncan Watts [37], with the *Small-World Network,* and Albert-László Barabasi [1-3], which developed the *Scale-Free Network.* About its work, *Barabási* has found that the websites that form the network (of the WWW) have certain mathematical properties. The conditions for these properties to occur are threefold. The first is that the network has to be expanding, growing. This precondition of growth is

[1] AMO - Advanced Modeling and Optimization. ISSN: 1841-4311

very important as the idea of emergence comes with it. It is constantly evolving and adapting. That condition exists markedly with the world wide web. The second is the condition of preferential attachment, that is, nodes (websites) will wish to link themselves to hubs (websites) with the most connections. The third condition is what is termed competitive fitness which in network terms means its rate of attraction.

Network Theory is an quickly expanding area of Network and Computer Sciences, and also may be considered a part of Graph Theory.

Complex Networks are everywhere. Many phenomena in nature can be modeled as a network, as brain structures, social interactions or the World Wide Web (WWW). All such systems can be represented in terms of nodes and edges. In Internet the nodes represent routers and the edges the 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.

These networks are not random. The topology of different networks are very close [31]. They follow from the Power Law, with a scale free structure. How can very different systems have the same underlying topological features? Searching the hidden laws of these networks, modeling and characterizing them are the current lines of research.

For their part, recall that according K. Mainzer [21-23],

"Symmetry and Complexity determines the spirit of nonlinear science" (2005).

And

"the universal evolution is caused by symmetry break, generating diversity, and increasing complexity and energy".

2. Graph Entropy

Graph theory has emerged as a primary tool for detecting numerous hidden structures in various information networks, including Internet graphs, social networks, biological networks, or more generally, any graph representing relations in massive data sets. Analyzing these structures is very useful to introduce concepts as Graph Entropy and Graph Symmetry.

We consider a functional on a graph, $G = (V, E)$, with P a probability distribution on its node (or vertex) set, V.

The mathematical construct called as *Graph Entropy* will be denoted by GE. It will be defined as

$$H(G,\ P) = min \sum\ p_i\ \log\ p_i$$

Observe that such function is convex.

It tends to $+\infty$ on the boundary of the non-negative orthant of R^n.

And monotonically to $-\infty$ along rays from the origin.

So, such minimum is always achieved and it will be finite.

The entropy of a system represent [28] the amount of uncertainty one observer has about the state of the system. The simplest example of a system will be a random variable, which can be shown by a node into the graph, being their edges the representation of the mutual relationship between them. Information measures the amount of correlation between two systems, and it reduces to a mere difference in entropies.

So, the *Entropy of a Graph* (from now, denoted by GE) 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 as the basic unit of Information is the bit, Entropy also may be viewed as the number of bits of "randomness" in the graph, verifying that

the higher the entropy, the more random is the graph

Let G be now an arbitrary finite rooted *Directed Acyclic Graph* (or *DAG,* in acronym). For each node, v, we denote $i(v)$ the number of edges that terminates at v. Then, the *Entropy of the graph* is expresable as

$$H\left(G\right)=\sum_{\substack{v\in V\\ i(v)\geq 2}}\left[i\left(v\right)-1\right]\ \log_2\left(\frac{Card(E)\ -\ Card(V)\ +\ 1}{i(v)-1}\right)$$

$H\left(X\right)$ may be interpreted in some different ways. For instance, given a random variable, X, it informs us about how random X is, how uncertainty we should about X, or how much variability X has.

3. Graph Symmetry

As we known, *Symmetry* into a system means invariance of its elements under a group of transformations [25]. When we take Network Structures [21, 22], it means invariance of adjacency of nodes under the permutations on node set.

Let G and H be two graphs. An *isomorphism* from G to H will be a bijection between the node sets of both graphs, i. e. a

$$f: G \to H$$

such that any two nodes, u and *v,* of G are adjacent in G iff *f(u)* and *f(v)* are also adjacent in *H.* Usually, it is called *"edge-preserving bijection".*

If an isomorphism exists between two graphs, G and *H,* then such graphs are called *Isomomorphic Graphs.*

The graph isomorphism is an equivalence, or equality, as relation on the set of graphs. Therefore, it partitions the class of all 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 an *isomorphism class of graphs.*

An *automorphism of a graph,* $G = (V,\ E)$, will be an isomorphism from G onto itself. So, a *graph-automorphism* of a simple graph, *G,* is simply a permutation on the set of its nodes, *V (G),*

$$f : G \to G$$

such that the image of any edge of G is always an edge in G. That is, if

$$e = \{u,\ v\} \in V\,(G)$$

then

$$f\,(e) = \{f\,(u)\,,\ f\,(v)\} \in V\,(G)$$

Either expressed in group theoretical way, we have

$$u \sim v \Leftrightarrow ug \sim vg \Leftrightarrow u^g \sim v^g$$

Being ug and vg (or also u^g and v^g, in the other very usual notation) the corresponding images of u and v under the permutation *g.*

The family of all automorphisms of a graph G is a permutation group on *V(G).* The inner operation of such group is the composition of permutations. Its name is very well-known, the *Automorphism Group of G,* and abridgedly, it is denoted by $Aut(G)$.

And conversely, all groups may be represented as the automorphism group of a connected graph.

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

About another interesting concept in Mathematics, the word *"genus"* has different, but very related, meanings. So, in Topology, it depends on to consider orientable or non-orientable surfaces.

In the case of *connected and orientable surfaces,* it is an integer that represents the maximum number of cuttings, along closed simple curves, without rendering the resultant manifold disconnected. For this reason, we may said that it is the number of *"handles"* on it. Usually, it is denoted by the letter *g.*

It will be also definable through the Euler number, or *Euler Characteristic,* denoted by χ.

Such relationship will be expressed, for *closed surfaces,* by

$$\chi = 2 - 2g$$

When the surface has b boundary components, this equation transforms to

$$\chi = 2 - 2g - b$$

which obviously generalizes the above equation.

For example, a *sphere,* an *annulus,* or a *disc* have genus $g = 0$.

Instead of this, a *torus* has $g = 1$.

In the case of *non-orientable surfaces,* the *genus* of a closed and connected surface is a positive integer, representing the number of cross-caps attached to a sphere.

Recall that a *cross-cap* is a two-dimensional surface that is topologically equivalent to a Möbius string.

As in the precedent analysis, it can be expressed in terms of the Euler characteristic, by

$$\chi = 2 - 2k$$

being k the *non-orientable genus.*

For example, a *projective plane* has non-orientable genus $k = 1$.

And a *Klein bottle* has a non-orientable genus $k = 2$.

Turning to graphs, its corresponding genus will be the minimal integer, *n,* such that the graph can be drawn without crossing itself on a sphere with n *handles.* So, a *planar graph* has genus $n = 0$, because it can be drawn on a sphere without self-crossing.

In the *non-orientable case,* the genus will be also the minimal integer, n, such that the graph can be drawn without crossing itself on a sphere with n *cross-caps.*

If we pass now to topological graph theory, we will define as *genus of a group, G,* the minimum genus of any of the undirected and connected Cayley graphs for G.

From the viewpoint of the Computational Complexity, *the problem of "graph genus" is NP-complete* [Thomassen, 1989].

We will says either *graph invariant* or *graph property,* when it depends only of 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 its possible isomorphisms of the graph. Therefore, it will be *a property of the graph itself,* not depending on the representation of the graph.

The semantic difference also consists in its character *quantitative* or *quantitative.*

For instance, when we said

"the graph does not possess directed edges"

this will be a *property,* because it is a *qualitative* statement.

While when we says

"the number of nodes of degree two in such graph"

this would be an *invariant,* because it is a *quantitative* statement.

From a mathematically strict viewpoint, *a graph property* can be interpreted as *a class of graphs,* composed by the graphs that have in common the accomplishment of some conditions.

Hence, also can be defined a graph property as a function whose domain would be the set of graphs, and its range will be the bivalued set composed by two options, true and false, $\{T,\ F\}$, according which a determinate condition is either verified or violated for the graph.

A graph property is called *hereditary*, if it is inherited by its induced subgraphs. And it is *additive,* if it is closed under disjoint union.

For example, the property of a graph to be planar is both additive and hereditary. Instead of this, the property of being connected is neither.

The computation of certain graph invariants may be very useful, with the purpose to 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 however, two graphs with the same invariants may or may not be isomorphic between them. So, we will arrive to the notion of *completeness.*

Let $I(G)$ *and* $I(H)$ be invariants of two graphs, G *and* H.

It will be considered *complete*, if the identity of the invariants ever implies the isomorphism of the corresponding graphs, i. e.

$$I(G) = I(H) \Rightarrow G \approx H$$

A *directed graph,* or *digraph,* is the usual pair $G = (V,\ E),$ buth now with an additional condition: it have at most one directed edge from node i to node j, being $1 \leq i,\ j \leq n$. We add the term "*acyclic*" when there are no cycles of any length. Usually, we use the acronym DAG to denote an acyclic directed graph.

A very important result may be this:

For each n, the cardinality of the $n - DAGs$, or *DAGs with* n *labeled nodes,* is equal to the number of $(n \times n) - matrices$ of 0′*s* and 1′*s* whose eigenvalues are positive real numbers, i. e.

$$\lambda_i \in \mathbf{R}_+, \forall i \in \{1, 2, ..., n\}$$

A previous result, due to Cvetkovic, Doob, and Sachs, said that
A digraph contains no cycle iff all eigenvalues of its adjacency matrix are equal to zero.

It is possible to prove that

Every group is the automorphism group of a graph.

If the group is finite, the graph may be taken to be finite.

And George Pólya observed that

Not every group is the automorphism group of a tree.

4. Complex Networks

We recall previously that given a random variable, X, its *Shannon Entropy* is given by

$$H(X) = -\sum P(x) \ \log_2 P(x)$$

whereas the *Rényi Entropy of order* $\alpha \neq 1$ of such random variable is

$$H_\alpha(X) = \frac{1}{1-\alpha} \log_2 \left(\sum P(x)^\alpha\right)$$

The Renyi Entropy of order α converges to the Shannon Entropy, when $\alpha \to 1$, i.e.

$$\lim_{\alpha \to 1} \left\{ \frac{1}{1-\alpha} \log_2 \left(\sum P(x)^\alpha\right) \right\} = -\sum P(x) \ \log_2 P(x)$$

So,

$$\lim_{\alpha \to 1} H_\alpha(X) = H(X)$$

Therefore, the Rényi Entropy may be considered as a generalization of the Shannon Entropy, or dually expressed, the Shannon Entropy will be a particular case of Rényi Entropy.

The structural information content will be the entropy of the underlying graph topology.

A method for determining the entropy of graphs, and therefore, of Complex networks, is possible, essentially due to [7]. In such procedure, we assign a probability value to each node. For these, we use an information functional which quantifies the structure. So, it allows us determining its entropy.

Firstly, it will be convenient to introduce a new geometrical tool, the so-called *j-spheres* of a graph.

Given an unlabeled and connected n-graph, G, and v_i one of its nodes. Then, the *j-sphere of* v_i is

$$S_j\left(v_i, G\right) = \{v \in V:\ d\left(v_i,\ v\right) = j,\ being\ j \geq 1\}$$

From these system of expanding-contracting circles and the cardinalities of their node set which each one of them contains, we introduce the information functional, denoted as f^V, by

$$f^V\left(v_i\right) = \exp_\beta\left\{k_1\ card\ \left[S_1\left(v_1, G\right)\right] +\ ...\ +\ k_n\ card\ \left[S_n\left(v_n, G\right)\right]\right\}$$

where $k_\tau > 0,\ 1 \leq \tau \leq n,\ \beta > 0.$

Their signification is that it shows the structural information, being the coefficients, β and k_τ, useful weighting the different characteristics of the graph in each j-sphere.

The *probability value associated to each node* is given by

$$P^V\left(v_i\right) = \frac{f^V\left(v_i\right)}{\sum_{j\ =1}^{card(V)} f^V\left(v_j\right)}$$

And so, the *Entropy of G* would be expressable as

$$H_f^V\left(G\right) = -\sum_{i=1}^{card(V)} P^V\left(v_i\right)\ \log P^V\left(v_i\right)$$

A network is said *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 have elements different to the identity.

Current research have revealed a very surprising result, according which the interaction networks displayed by most complex systems are highly heterogeneous [31].

5. Conclusions

Statistical entropy is a probabilistic measure of uncertainty [11, 35], or ignorance about data. Whereas *Information* is a measure of a reduction in that uncertainty [6, 17]. Whereas the Entropy of a probability distribution is just the expected value of the information of such distribution [26].

All these improved tools must permits to advance not only in fields as Optimization Theory, but also on Generalized Fuzzy Measures [12-14, 36], Economics [15], modeling Biological Systems, for Robustness of the processes, Information Difussion, Synchronization, and so on.

References

[1] Albert, Jeong, and Barabási: "The Diameter of the WWW". *Nature,* Vol. 401, pp. 130-131, 1999.

[2] Barabási: *Linked: How Everything Is Connected to Everything Else.* Plume Publ., 2002.

[3] Barabási, Newman, and Watts: *The Structure and Dynamics of Networks.* Princeton University Press, 2006.

[4] Burbea, and Rao: "Entropy on Intuitionistic Fuzzy Sets and Interval-Valued Fuzzy Sets". *FSS,* Vol. 78, pp. 305-316, 1996.

[5] Chaitin: "Algorithmic Entropy of Sets". *Comp. and Maths with Appls.* Vol. 2, pp. 233-245, 1976.

[6] Devine: " The Insights of Algorithmic Entropy". *Entropy* journal, Vol.11, Issue No. 1, pp. 85-110. MDPI Publishers, Basel, 2009.

[7] Dehmer: "Information processing in Complex Networks: Graph entropy and Information functionals". *Appl. Maths. and Comp.* Vol. 201, Issues 1-2, pp. 82-94, 2008.

[8] Dumitrescu: Entropy of a fuzzy dynamical system, FSS, Vol. 70, pp. 45-57, 1995.

[9] Dumitrescu: Entropy of a fuzzy process. FSS. Vol. 55, pp. 169-177, 1993.

[10] Dumitrescu: Fuzzy measures and the entropy of fuzzy partitions. J. Math. Anal. Appl. Vol. 175, pp. 359-373, 1993.

[11] Garmendia: "The Evolution of the Concept of Fuzzy Measure". *Studies in Computational Intelligence,* Vol. 5, pp. 186-200, 2005.

[12] Garrido: "Additivity and Monotonicity in Fuzzy Measures". Plenary Talk in *ICMI45,* at Bacau University, then published at *SCSSM journal,* Vol. 16, pp. 445-458, 2006.

[13] Garrido: "Information and Entropy Measures". Plenary Talk in the *6th ICTAMI,* at Alba Iulia University, accepted; must to be published at *AUA journal,* Vol. 19 (special issue), pp. 18, 2009.

[14] Garrido: "Analyzing the concept of Entropy". *EIJ-AMO* journal, Vol. 11, Issue No. 4, pp. 525-530, 2009.

[15] Georgescu-Roegen: *The Entropy Law and the Economic Process.* Harvard University Press. Cambridge, MA,1971.

[16] Hung, and Yang: "Fuzzy entropy on Intuitionistic Fuzzy Sets". *Intern. J. of Intelligent Systems.* Vol. 21, pp. 443-451, 2006.

[17] Jaynes: "Information Theory and Statistical Mechanics". *Phys. Rev.* Vol. 108, pp. 171-190, 1956.

[18] Jozsa: "Quantum Information and Its Properties". Chapter in the book *Introduction to Quantum Computation and Information.* Eds. Lo, Popescu, and Spiller. World Scientific, 1998.

[19] Liu: "Entropy, distance and similarity measures of fuzzy sets and their relations. *FSS,* Vol. 52, pp. 305-318, 1992.

[20] Luca, and Termini: "A definition of non-probabilistic entropy, in the setting of fuzzy theory. Information and Control. Vol. 20, pp. 301-318, 1972.

[21] Mainzer: *Symmetry and Complexity: The Spirit and Beauty of Nonlinear Science.* World ScientificSeries on Nonlinear Science, Series A, Vol. 51.

World Scientific Publ.Co., 2005.

[22] Mainzer: Symmetries of Nature: *A Handbook for Philosophy of Nature and Science.* Walter de Gruyter, 1996.

[23] Mainzer: *Thinking in Complexity: The Computational Dynamics of Matter, Mind, and Mankind.* Springer, 2007.

[24] Passarini et al.: "The von Neumann entropy of networks". *MPRA,* Paper No. 12538. University of Munich, 2009.

[25] Neumann: *Grundlagen der Quantenmechanik.* Springer Verlag, Berlin, 1955.

[26] Rényi: "On measures of information and entropy". *Proc. of the 4th Berkeley Symposium on Mathematics, Statistics and Probability,* pp. 547-561, 1961.

[27] Shannon: "A Mathematical Theory of Communication", Bell System Technical Journal, pp. 379-423 and 623-656, 1948. Later, it appears as book, at Illinois Press, 1963. Also by E-print.

[28] Simonyi: "Graph Entropy: A Survey". *DIMACS Series in Discrete Mathematics and Theoretical Computer Science.*

[29] Sinai: "On the concept of Entropy of a Dynamical System". *Doklady Akademii Nauk SSSR,* Vol. 124, pp. 768-771, 1959.

[30] Sinai: "On the foundations of the Ergodic Hypothesis for a Dynamical System of Statistical Mechanics". *Doklady Akademii Nauk SSSR,* Vol. 153, pp. 1261-1264, 1963.

[31] Strogatz: "Exploring Complex Networks". *Nature,* Vol. 410, pp. 268-276, 2001.

[32] Titchener, et al.: "Deterministic Complexity and Entropy". *Fundamenta Informaticae,* Vol. 64, Issue 1-4, Contagious Creativity-In Honor of the 80th Birthday of Prof. Solomon Marcus, 2004.

[33] Titchener: "A measure of information". *Proc. Data Compression Conference,* pp. 353-362, Snowbird, UT, 2000.

[34] Titchener:"A deterministic theory of Complexity, Information and Entropy". *IEEE Information Theory Workshop.* San Diego, CA, 1998.

[35] Volkenstein: *Entropy and Information.* Series: Progress in Mathematical Physics, Vol. 57, Birkhäuser Verlag, 2009.

[36] Wang, and Klir: *Generalized Measure Theory.* Springer Verlag, Berlin-New York, 2008.

[37] Watts: *Small Worlds: The Dynamic of Networks between Order and Randomness.* Princeton Studies in Complexity. Princeton University Press, 2003.

[38] Wehrl: "General properties of entropy". *Rev. Mod. Phys.* Vol. 50, pp. 221-260, 1978.

[39] Wilson: *Consilience.* Knopf Publ. New York, 1998.

[40] You, and Gao: "Maximum entropy membership functions for discrete fuzzy variables". *Information Sciences,* 179, Issue No. 14, pp. 2353-2361, 2009.