Entropy of Various Matrix Energies for Complex Networks

Mikołaj Morzy[1] & Tomasz Kajdanowicz[2]

[1]Poznań University of Technology, Poland, Mikolaj.Morzy@put.poznan.pl

[2]Wrocław University of Science and Technology, Poland, Tomasz.Kajdanowicz@pwr.edu.pl

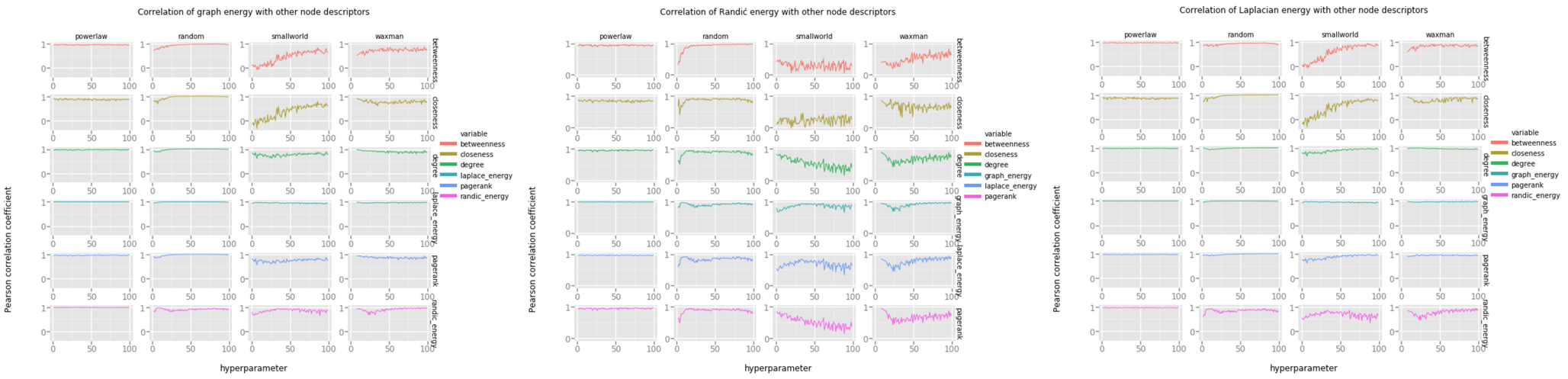

Figure 1: Correlation of graph energies with node descriptors

Abstract

Matrix energy is the sum of the absolute values of its eigenvalues. In the context of networks, several different energies have been proposed, depending on the particular form of the network matrix. Examples include the regular energy defined over the adjacency matrix, the Laplacian energy defined over the Laplacian of the network, or the Randić energy based on the so-called Randić matrix, in which each entry is the inverse of the square root of the product of node degrees, if the two nodes are adjacent. The main problem with the above definition of network energy is the fact that it is hardly operational. Defined over the entire network matrix, the energy of that matrix is a single number trying to describe the network. In large complex networks, there can exist heterogeneous parts of the network with highly diversified topologies and properties, resulting in significant variation of local energy.

Introduction

In this work, we diverge from the traditional approach of computer science, where entropy is understood in the light of Shannon's information theory. Instead, we use entropy in the physical sense, as the measure of the dispersion of energy, where energy is computed from various network-invariant matrix representations. We look at how the entropy of various network energies changes with the gradual change of network topology and we compare energies of node egocentric networks with other well-known node descriptors, such as in-degree, out-degree, betweenness, local clustering coefficient, and other indexes. The energies computed in this experiment include graph energy, Randić energy, and Laplacian energy. Each of these energies describes a particular matrix, which in turn, expresses a different aspect of a graph. Below we present a brief overview of each energy.

Graph energy of a graph is defined as $E_G(G) = \sum_{i=1}^{n} |\mu_i|$, where $\mu_1, \ldots, \mu_n$ are the eigenvalues of the adjacency matrix M_A (also known as the *spectrum* of the graph).

Randić matrix of the graph $G = \langle V, E \rangle$ is defined as:

$$M_R(i,j) = \begin{cases} 0 & if \quad i = j \\ \frac{1}{\sqrt{d_i d_j}} & if \quad (i,j) \in E \\ 0 & if \quad (i,j) \notin E \end{cases}$$

Randić energy of a graph is defined as $E_R(G) = \sum_{i=1}^{n} |\rho_i|$, where $\rho_1, \ldots, \rho_n$ are the eigenvalues of the Randić matrix M_R.

The Laplacian matrix of the graph $G = \langle V, E \rangle$ is defined as:

$$M_L(i,j) = \begin{cases} d_i & if \quad i = j \\ -1 & if \quad i \neq j \ \& \ (i,j) \in E \\ 0 & otherwise \end{cases}$$

Laplacian energy of a graph is defined as $E_L(G) = \sum_{i=1}^{n} |\lambda_i - \frac{2m}{n}|$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of the Laplacian matrix M_L, $m = |E|$ is the number of edges, and $n = |V|$ is the number of nodes.

One can compute the energy of the entire graph, but this is of little use, since major variations may occur between different parts of the graph. In the experiment we compute local energies (energies of ego-networks for each node), thus obtaining a random variable representing all energy 'states' of the graph. We then turn these into probability distributions and measure the entropy of each energy distribution. The assumption is that the entropy of energy distribution may reveal to which extent the energy is uniformly distributed across the network.

In the experiments we use for artificial network models which represent major topological families of networks: Erdos-Renyi random model, Watts-Strogatz small world model, Waxman model, and Holme-Kim powerlaw model. For each network model we generate several instances of the model, modifying a single graph parameter. Then, we compute all the energies of all ego-networks (of radius $r = 1$) and we compute the entropy of the distribution of each energy.

The gradual changes of each model are realized by modifying the following parameters:

- **Erdos-Renyi** random model: the probability of creating an edge between a random pair of vertices changes uniformly from $p = 0.01$ to $p = 1.0$, the network changes gradually from a set of isolated nodes to a clique
- **Watts-Strogatz** small world model: the probability of randomly rewiring an edge changes uniformly from $p = 0.01$ to $p = 1.0$, the network changes gradually from a strictly ordered structure where each node links to its $k = 4$ nearest neighbors with no long-distance bridges to a fully random network
- **Waxman** random geometric model: the Waxman models place n nodes uniformly at random in a rectangular domain, two nodes u, v are connected with an edge with probability $p = \alpha * e^{\frac{-d}{\beta * L}}$, the distance d is the Euclidean distance between the nodes u and v, L is the maximum distance between all nodes in the graph. In this experiment we change α uniformly in the range $[0.01, 1.0]$ and $\beta = 0.1$.
- **Holme-Kim** powerlaw model: the probability of closing a triangle after adding a random edge changes uniformly from $p = 0.01$ to $p = 1.0$, the resulting network is basically the same as the Albert-Barabasi model with an added triangle closure step which increases the local clustering coefficient.

Main Objectives

1. Investigate properties of energies of various matrices representing complex networks
2. Measure the correlation between energies of ego-networks and other vertex descriptors
3. Compute the entropy of energies for different network topologies

Results

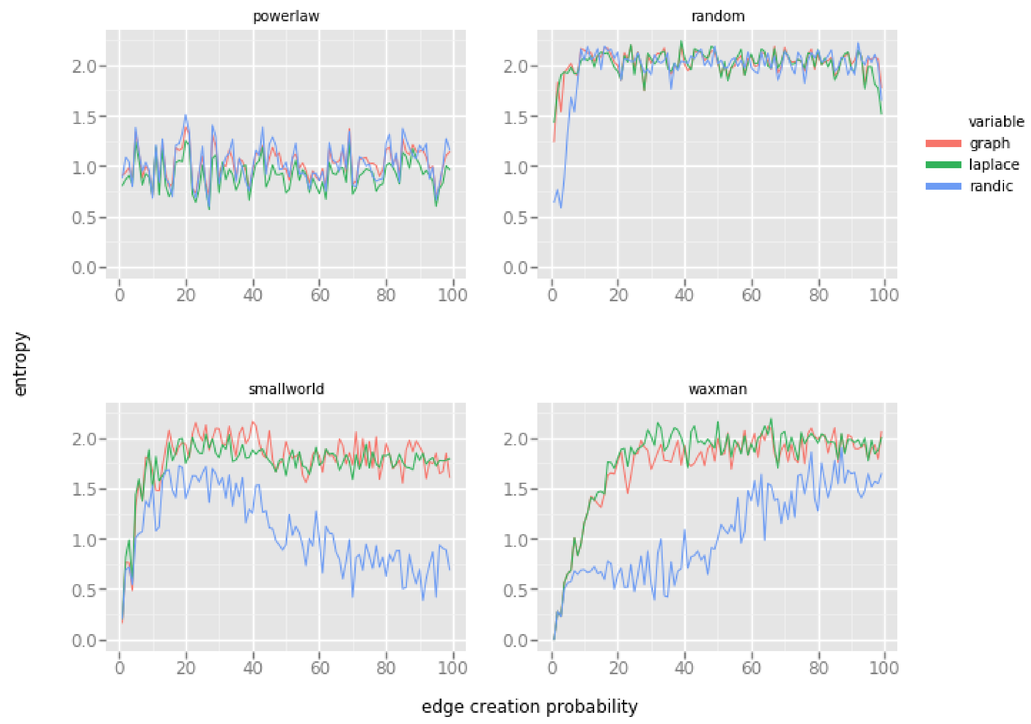

Figure 2: Entropy of energies for different network topologies

Conclusions

- **Erdos-Renyi** random model : entropies of all energies quickly increase and stay at the maximum level during the densification of the graph, when the edge probability creation reaches 1 (leading to a large clique), the entropies drop.
  - Graph energy correlates strongly with closeness and pagerank, even for relatively low probability of random edge creation (20%), in these networks there is enough connectivity between nodes to produce large connected components, but the networks are far from being cliques
  - Randić energy correlates well with the betweenness, one can investigate the possibility of estimating the betweenness of a node based on its Randić energy for a wide spectrum of random networks
  - Laplacian energy correlates almost perfectly with degree, pagerank and closeness, since both pagerank and closeness are costly to compute, one can estimate these values based on the Laplacian energy of a node
- **Watts-Strogatz** small world model: we begin with a regular small world lattice and gradually increase the probability of edge rewiring. Interestingly, the addition of random edges affects the entropies only at the beginning, but after a certain threshold the entropy begins to diminish
  - Graph energy correlates only with Laplacian energy, there is some correlation with degree and pagerank, but most probably this correlation is too weak to provide accurate estimates
  - Randić energy cannot be reasonably used to provide any estimates regarding the properties of nodes in small world networks
  - Laplacian energy correlates to some extent with degree and pagerank, providing means for estimation, but this can be achieved only when the probability of edge rewiring is significant
- **Waxman** random geometric model: the behavior of entropies is very similar to the small world network, despite the fact that the generative model is significantly different.
  - Graph energy correlates with degree and pagerank for small values of $alpha$ and $beta$ parameters of the model
  - Randić energy again does not seem to be very useful for estimating the values of other node descriptors
  - Laplacian energy correlates very strongly with degree, and somehow strongly with pagerank
- **Holme-Kim** powerlaw model : when the probability of triad closure in the powerlaw graph model is low, the entropy tends to increase, but after reaching a certain saturation point the entropy of most energies begins to steadily decrease
  - Graph energy correlates with betweenness, degree and powerlaw to the extent that allows us to suspect that it is possible to estimate these descriptors based on the graph energy of the node
  - Randić energy correlates well with other types of node energies, but the correlations with degree, pagerank and betweenness are significant and can be useful
  - Laplacian energy exhibits very strong correlation with degree, betweenness and pagerank