Zipf's Law in the Dynamical Importance of Network Nodes and Links

Jing Ma
Physics Department, Boston University

December 20, 2016

Abstract

In network science, dynamical importance can be used to quantify the importance of nodes and links. In physical and social sciences, Zipf's law is found to be able to describe many types of data. In this article, the distributions of dynamical importance of nodes and links will be analyzed for several types of undirected networks, i.e. regular network, Erdos Renyi random network, and Barabasi Albert scale free network. We will show that the distributions of dynamical importance of nodes and links also obey Zipf's law.

1 Background

We all know the fact that in a network, some nodes and links are important, while others are not. It's then essential to find a quantity to quantify the importance of nodes and links in a network. It's tempting to choose node degree as the quantity we want, since the most basic property of a node is the degree, and the type of a network is always defined by its degree distribution. However, a higher degree doesn't always mean a higher importance. For example, in a disassortative network, nodes with high degrees are important, as expected; but nodes with low degrees can play a significant role in the network as well, as they tend to be the only bridges among communities. Besides, degree is a local property, and hence can't reflect the structure of the whole network.

A quantity being able to measure the real importance of nodes and links is needed, and it should reflect the structure of the whole network. Based on the requirements, a quantity called dynamical importance was introduced [4]. It has been discussed in many papers that the largest eigenvalue of the network adjacency matrix, which we call λ, is always a determinant in the properties of different dynamical networks [3, 5, 1, 2]. The dynamical importance is then defined as the percentage decrease in λ upon the removal of the node or the link, i.e. for a node, $I_k \equiv -\frac{\Delta\lambda_k}{\lambda}$, where k is the label of the node; for a link, $I_{i,j} \equiv -\frac{\Delta\lambda_{i,j}}{\lambda}$, where i, j are the labels of the nodes connected by the link.

2 Simulations of Different Networks

As discussed above, the importance that a node or a link plays could be different, depending on the structure of the network. What we are interested in here is the distributions of the dynamical importance of nodes and links. Simulations are done based on four types of network models defined in the Python library named 'networkx', i.e. regular graph, Erdos Renyi random graph, Watts Strogatz small world graph, and Barabasi Albert scale free graph. For each type of network, 1000 nodes are generated and connected according to its corresponding dynamic properties. The dynamical importance I_k for each node and $I_{i,j}$ for each link are calculated, and then sorted in descending order.

3 Results of Different Networks

3.1 Regular Graph

For a regular graph, the dynamical importance distributions are shown in Figure 1. We can see that all of the nodes or links are of almost the same importance, which is consistent with our intuition since no nodes or links have privileges over others in a regular graph. And the distributions are also plotted in a log-log scale in Figure 2 for consistency with the following parts. A linear regression is done to the 20% of nodes or links with larger importance and the slope is around zero, as expected.

3.2 Erdos Renyi Random Graph

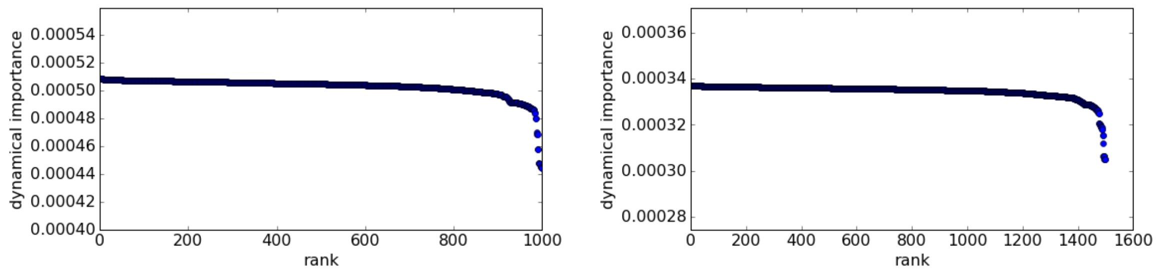
For a random graph, the dynamical importance distributions are shown in Figure 3. They look like power laws in the linear scale graphs so the distributions are then plotted in a log-log scale in Figure 4. Similarly as above, a linear regression is done to the 20% of nodes or links with larger importance. We can see that the distributions indeed behave linearly in the log-log scale graphs, which means they obey power laws themselves. The regression slope, and thus the power law exponent, is -0.33 for nodes and -0.30 for links.

3.3 Watts Strogatz Small World Graph

For a small world graph, the program calculating the largest eigenvalue λ doesn't always converge when certain nodes or links are removed. We will just skip this type of network.

3.4 Barabasi Albert Scale Free Graph

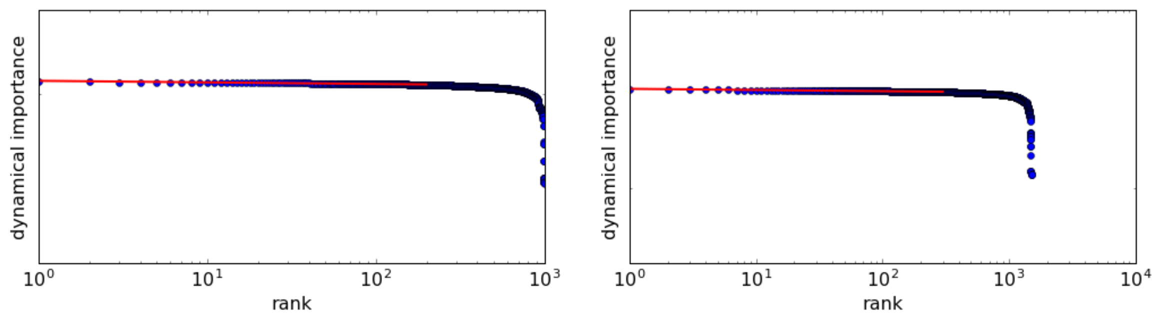
For a scale free graph, the dynamical importance distributions are shown in Figure 5. Similarly as above, they are plotted in a log-log scale in Figure 6, and a linear regression is done to the 20% of nodes or links with larger importance. We can still find power laws in the distributions of dynamical importance of nodes and links. The regression slope, and thus the power law exponent again, is -0.80 for nodes and -0.78 for links.



(a) nodes

(b) links

Figure 1: the dynamical importance distributions for a regular graph in a linear scale



(a) nodes

(b) links

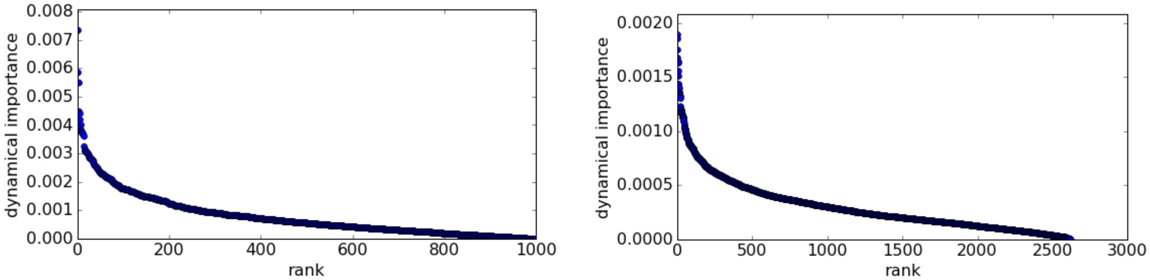
Figure 2: the dynamical importance distributions for a regular graph in a log-log scale

4 Discussion

As Zipf's law says, "many types of data studied in the physical and social sciences can be approximated with a Zipfian distribution, one of a family of related discrete power law probability distributions." [6] To be brief, it means some quantity is in a power law in terms of its rank. It's clear from our simulations that the distributions of dynamical importance of nodes and links in a network obey perfect Zipf's law. Zipf's law is not only used to describe the behavior of natural languages, but also valid in other physical and social fields.

5 Conclusion and Outlooks

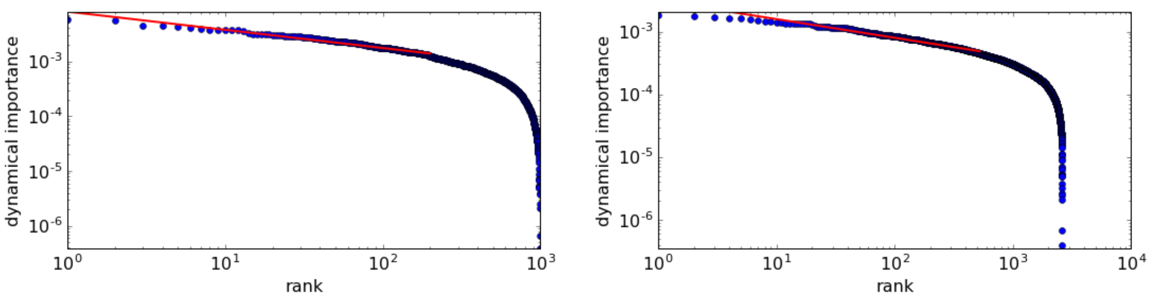
In this article, we show that for many different types of networks, i.e. regular network, Erdos Renyi random network, and Barabasi Albert scale free network, the distributions of dynamical importance of nodes and links obey Zipf's law. Further efforts will be made so that more types of networks will be investigated and importance based on other definitions can be compared, which we might address in a future publication.



(a) nodes

(b) links

Figure 3: the dynamical importance distributions for a random graph in a linear scale



(a) nodes

(b) links

Figure 4: the dynamical importance distributions for a random graph in a log-log scale

Acknowledgements

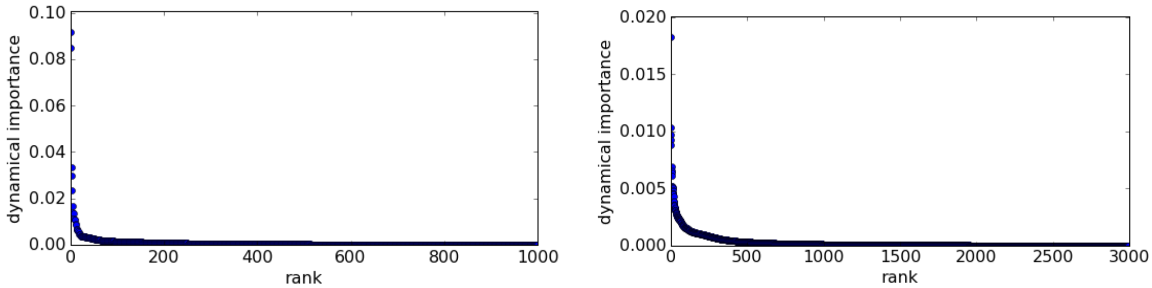
This article is a project article of the course PY 895 Network Science from Boston University. We thank Prof. Eugene Stanley who provided this course, brought cutting-edge knowledge into the class, and encouraged us to come up with our new ideas. We also thank Dr. Jianxi Gao, who was the TA of this course, for the inspiring comments and suggestions.

References

[1] nD Cvetković and Peter Rowlinson. The largest eigenvalue of a graph: A survey. *Linear and multilinear algebra*, 28(1-2):3–33, 1990.

[2] Charles R MacCluer. The many proofs and applications of perron's theorem. *Siam Review*, 42(3):487–498, 2000.

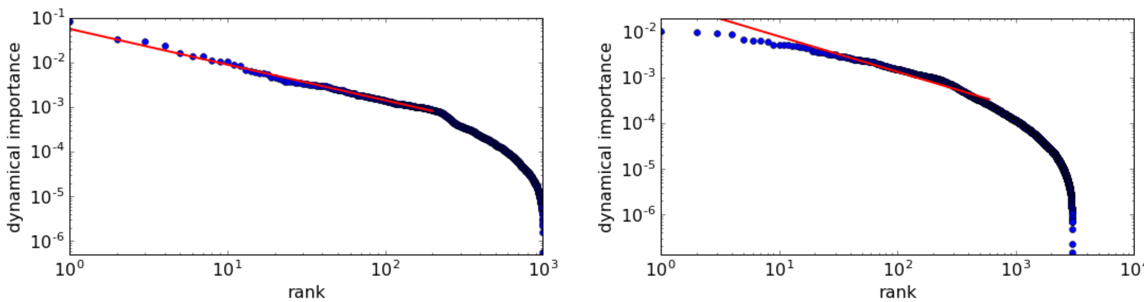
[3] Juan G Restrepo, Edward Ott, and Brian R Hunt. Onset of synchronization in large networks of coupled oscillators. *Physical Review E*, 71(3):036151, 2005.



(a) nodes

(b) links

Figure 5: the dynamical importance distributions for a scale free graph in a linear scale



(a) nodes

(b) links

Figure 6: the dynamical importance distributions for a scale free graph in a log-log scale

[4] Juan G Restrepo, Edward Ott, and Brian R Hunt. Characterizing the dynamical importance of network nodes and links. *Physical review letters*, 97(9):094102, 2006.

[5] Yang Wang, Deepayan Chakrabarti, Chenxi Wang, and Christos Faloutsos. Epidemic spreading in real networks: An eigenvalue viewpoint. In *Reliable Distributed Systems, 2003. Proceedings. 22nd International Symposium on*, pages 25–34. IEEE, 2003.

[6] Eric W. Weisstein. Zipf's law. *MathWorld*.