

Single Node In-Out Degree Correlations

Description

One-point correlations are the correlations between indegree and outdegree of a node of a directed network. Nodes with large indegree are *hubs*, nodes with large outdegree are *authorities*. To calculate the correlation between indegree and outdegree is a way to check whether hubs are also authorities or not.

Pros & Cons

The network to analyze must be directed, otherwise there are no special constraints.

Applications

Useful in many contexts. As an example, for the World Wide Web, the study of one-point correlations allows to address the question: are the most linked URLs also the ones that link more pages?

Implementation Details

The algorithm requires two inputs, the file where the edges of the network are listed and the number of points one wishes to have in the binned correlation function described below. A first read-in of the inputfile will set the values of the number of nodes and edges of the network. In the second read-in the indegrees and outdegrees of all nodes will be calculated and the edges are stored in an array. The algorithm calculates first the average product $\langle k_{in} k_{out} \rangle$ of indegree and outdegree for the same node. In the absence of correlations between indegree and outdegree

$$\langle k_{in} k_{out} \rangle = \langle k_{in} \rangle \langle k_{out} \rangle,$$

so the crossed correlation coefficient

$$\frac{\langle k_{in} k_{out} \rangle}{\langle k_{in} \rangle \langle k_{out} \rangle} = 1$$

on an uncorrelated network. If the crossed correlation coefficient is sensibly different from one, one-point correlations are important. The program displays the crossed correlation coefficient in the NWB console. In addition, the application generates two output files, corresponding to two different ways of partitioning the interval spanned by the values of indegree. In the first output, the outdegree is averaged among all nodes with equal indegree; the result is displayed in the output together with the corresponding indegree. The second output gives the binned correlation function, i.e. the interval spanned by the values of indegree is divided into bins whose size grows while going to higher values of the variable. The size of each bin is obtained by multiplying by a fixed number the size of the previous bin. The program calculates the average outdegree for all nodes whose indegree falls within each bin. This technique is particularly suitable to study skewed correlation functions: the fact that the size of the bins grows large for large indegree values compensates for the fact that not many nodes have high indegree values, so it suppresses the fluctuations that one would observe by using bins of equal size. On a double logarithmic scale the points of the correlation function will appear equally spaced on the x-axis. The program runs in a time $O(m)$, m being the number of edges of the network.

Acknowledgements

The algorithm was implemented and documented by S. Fortunato, integrated by S. Fortunato and W. Huang.

References

Albert, R., and Barabasi, A.-L. (2002) [Statistical mechanics of complex networks](#). Review of Modern Physics 74:47-97.

Newman, M.E.J. (2003) [The structure and function of complex networks](#). SIAM Review 45:167-256.

Pastor-Satorras, R., Vespignani, A.(2002) Evolution and Structure of the Internet. Cambridge University Press.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U.(2006) [Complex networks: Structure and dynamics](#). Physics Reports 424: 175-308.