Networks in Economics and Social Science
2019/2020
Homework 4

Roberto Casarin

Abstract

The goal of this homework is to study which random graph model among Erods-Renyi (ER), Watts-Strogatz (WS) and Barabasi-Albert (BA), is the best theoretical model for the network data provided. Your report should provide a reply to the questions given in this document. You can use the LaTeXsource file of this document to generate your report. Your Homework is due by 13th of April by e-mail (`r.casarin@unive.it`).

1 Load, visualize and analyse network data

1. Load the interlocking directorate (ID) data (nodes are companies, edges are common board members) from the MATLAB workspace $VIPD.mat$ and use the first 2 columns of the edge list `VIPDedgelist` to generate a MATLAB graph object (see video lectures and background material).

2. Generate a subgraph by removing the isolated nodes (Hint: evaluate the node centrality `d` and use the function `subgraph(Ginitial,d>=1)`)

3. Exhibit a plot of the graph, use the layout `force` (see Class Videos).

4. Exhibit the degree distribution of the new graph.

5. Find the average degree.

6. Find the average clustering coefficient.

7. Find the average path length.

8. Find the number of nodes of the graph (Hint: use the properties of MATLAB graph object created)

2 Calibrate random network models

Consider the following models presented in our classes:

- The ER model $G_{ER}(n, p)$ where n is the number of nodes and p the edge probability.
- The WS model $G_{WS}(n, K, p)$ where n is the number of nodes, $2K$ the number of neighbours of a given node in the ring lattice regular graph, and p is the re-wiring probability.
- The BA model $G_{BA}(n, m, m_0)$ with n the number of nodes, m the initial number of connected nodes, and m_0 the number of edges generated at each iteration of the preferential attachment mechanism.

1. For the ER, WS, and BA models, find the value of the parameters which make the average node degree of each model close to the one of the interlocking directorate (ID) network.
2. Provide in a table the values of: the average clustering coefficients, the average path length and the average degree; for the three calibrated models and the ID data.
3. Provide the histogram of the degree distribution for models and data in the same plot.
4. Conclude, explaining which is the best model for the ID network.