Networks in Economics and Social Science
Homework 2019/2020

Roberto Casarin

Abstract

Prepare a report containing the analysis of the interlocking 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 16th of March by e-mail (`r.casarin@unive.it`).

1 Dataset description

The nodes of the network are companies. If the same individual is in the boards of two companies then there is an edge between those two companies. This set of nodes and edges defines an *interlocking directorate network*.

The two files `VIPDedgelist.csv` and `VIPDnodelist.csv` contain the edge and node list of the interlocking directorate network for two Italian provinces (Vicenza and Padova).

The vertex set, which comprises 33340 nodes, decomposes in two subsets following the province (see column partition in the node list file).

Multiple edge between nodes are allowed. There are three types of edges: Female, Male and Firm, which correspond to three different type of board members (directors) in the interlocking directorate.

The edge set, which includes 16896 edges, decomposes in three subsets following the type of edge (see column partition in the edge list file, where 1 means Male, 2 Female and 3 Firm)

2 Giant component

Filter out the Giant component and apply Force Atlas 2 layout with parameters: scaling 2.0, gravity 10.0. Exhibit the Giant component (export the graph from the preview window to a `.png` format). Your network should be similar to the one in Figure 1. In your figure use **coloured nodes** following the node-partition

Keep the Giant-component filter running and create a new workspace with the subgraph given by nodes and edges of the Giant component. Work in the new workspace and provide:

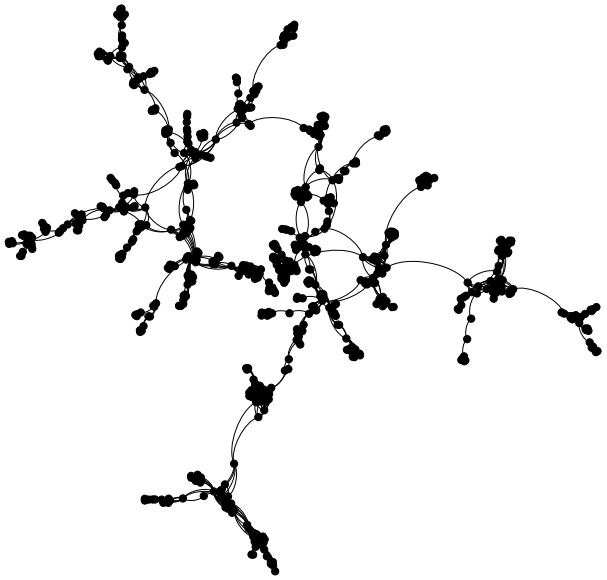

Figure 1: Giant component

1. The number of nodes and edges in the Giant component.
2. Find the average degree.
3. Find the average path length.
4. The maximum degree and the ID (from the Data Laboratory) of the node with the largest degree.
5. Put in red color the depth-2 ego network of the node with maximum degree and exhibit the graph.

Export the node list from the Data Laboratory session and provide the scatter plots for the following pairs of variables

1. Degree versus eigenvector centrality.
2. Degree versus closeness centrality.
3. Degree versus eccentricity.

Exhibit different pictures of the giant component (export them from the preview section) with **node size** (size between 1 and 10) depending on:

1. Node degree.
2. Eigenvector centrality.
3. Closeness centrality.
4. Betweenness centrality.
5. Eccentricity.

3 Bipartite graph

Consider the original network with 33340 nodes and 16896 edges. Remove the nodes with null degree, remove the edges between companies of the same province and keep in the graph only edges between companies of different provinces (hint: use intra-edges filter).

Exhibit the resulting bi-partite graph, with the node of one group (e.g., node-partition variable equal to 1) on the left-hand-side of the graph and nodes of the other group (e.g., node-partition variable equal to 2) on the right.

4 Betweenness centrality

The betweenness centrality indicates how relevant is a node in terms of connecting other nodes in the graph. Let $n(u, v)$ be the number of shortest paths P^*_{uv} from u to v, and $n_w(u, v) = |\{P^*_{uv}; w \in P^*_{uv}\}|$, i.e. the number of shortest paths from u to v going through the node w, then the betweenness centrality

$$\mathrm{Bet}_w(G) = \sum_{u \neq v, w \notin \{u,v\}} \frac{n_w(u, v)/n(u, v)}{(n-1)(n-2)} \tag{1}$$

where n is the number of nodes in the network.

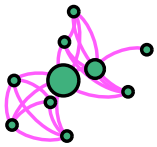

Figure 2: Giant component

Consider the original network with 33340 nodes and 16896 edges. Remove the nodes with null degree and the edges with edge-partition value equal to 1 (i.e. male directors). Export the graph to a new workspace.

How many nodes and edges can you find in the new subgraph?

Exhibit the resulting bi-partite graph with the nodes with node-partition value equal to 1 on the left-hand side, and the nodes with node-partition value equal to 2 on the right-hand side of the figure (hint: use filter to identify a partition, and the select tool to drag and drop them toward the left or the right).

Identify the Id (from the Data Laboratory) of the node with the largest betweenness centrality, and show that its max-depth ego network is equivalent to the one in Fig. 2.