Network 2

Giorgio Ricchiuti
www.grarchive.net
giorgio.ricchiuti@unifi.it

UNIVERSITÀ DEGLI STUDI FIRENZE

UNIVERSITÀ CATTOLICA del Sacro Cuore
CLE
Complexity Lab in Economics

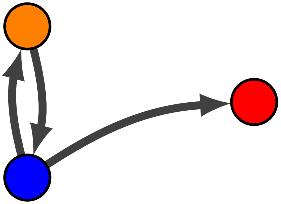
CompEc

A network is represented from a mathematical point of view by an **adjacency matrix**. The

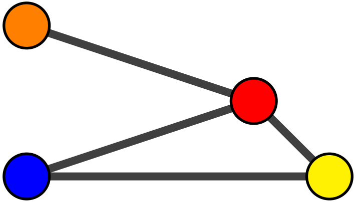
element of the **adjacency matrix** a_{ij} indicates that a link exists between nodes i and j; that is $a_{ij} = 1$ if investor i goes to country j; otherwise $a_{ij} = 0$. The **degree** of a node i is the number

of its links and is calculated by

$$k_i = \sum_j a_{i,j} \tag{1}$$



Direct Network



Indirect Network

Given the Topology we can calculate local, meso and large scal network topological measurements.

A) Local:
- at first order: degree;
- at second order: average neighbors degree;
- at third order: clustering;
- at fourth order: squared clustering.

B) Higher order measurements:
- betweenness;
- closeness;
- eigenvector centrality;
- eccentricity.

- The **degree** of a node i is the number of its links;

$$k_i = \sum_j a_{ij}$$

- the **degree centrality** is achieved dividing the degree by the number of nodes of the network:

$$dc_i = k_i/(N-1); \quad (2)$$

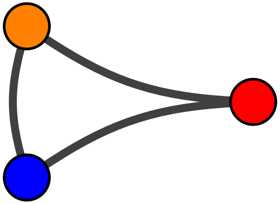
where N is the total number of nodes.

Starting from the concept of degree, we can calculate **the average neighbor degree** (Ave Neigh Degree), which generally depends on the degree of the node considered (Caldarelli, 2007).

Having a high *average* means that the node is linked to highly-connected nodes. This indicator can also help us to see if a network is **assortative**, i.e. if the nodes are similar in some way. To be more specific, if the average degree of the neighbors increases with the node's degree, then the network is assortative; if it decreases, then the network is disassortative.

Another local measure is the **clustering coefficient**, which is a measure of the density of connections around a vertex.

It enables us to calculate the proportion of the neighbors closest to the node that are connected to one another.



The **squared clustering** is the density of squared around the node. This measure captures the connections of the prime nodes that are neighbors of the node being analyzed.

A higher squared clustering coefficient means that the nodes are very connected to each other through another (common) node. The concept of density is reinforced and increases the possibility of contagion between the node and its neighbors.

The **distance** d_{ij} between two vertices i, j is *the shortest* number of edges to go from i to j. The neighbors of a vertex i are all the vertices j which are connected to that vertex by a single edge ($d_{ij} = 1$). Using the adjacency matrix this can be written as:

$$d_{ij} = min \sum a_{kl}$$

The **betweenness centrality** is based on dynamical properties of the graph and is given by the number of times that one vertex k is crossed by minimal path from one vertex i to j.

$$b_i = \sum_{\substack{j,l=1,N \\ i \neq j \neq l}} \frac{d_{jl}(i)}{d_{jl}} \qquad (3)$$

where d_{jl} is the total number of different shortest paths (distances) going from j to l and $d_{jl}(i)$ is the subset of those distances passing through i. The sum runs over all pairs with $i \neq j \neq l$.

The **closeness centrality** is:

$$cl_i = \frac{N-1}{\sum_j d_{ij}} = \frac{1}{\bar{d}_i} \tag{4}$$

which is the reciprocal of the average distance of one node from the other nodes. In order to be a hub, a node should not be very distant from all the others.

The **eigenvector centrality** measures the importance of a node based on the score of its neighbors. Contrary to the previous measures, this one is not based on distance among nodes, but depends recursively on the centrality of the neighbors. In vector notation, the eigenvector centrality c is the vector that solves the equation $W \cdot c = \lambda c$ where λ is the largest eigenvalue. In terms of recursive expression we can define the eigenvector centrality of node i as

$$c_i = \frac{1}{\lambda} \sum_j W_{ij} c_j \tag{5}$$

It is intrinsically based on the spectral properties of adjacency matrix. So it provides a different approach to assess node centrality.

The **eccentricity** of a node is

$$e_i = \frac{1}{\max_{\forall j \in N} d_{ij}} \tag{6}$$

that is the inverse of the maximum distance of that node from any other possible node of the network.

The **Katz centrality or alpha centrality** computes the relative influence of a node within a network by measuring the number of the immediate neighbors (first degree nodes) and also all other nodes in the network that connect to the node under consideration through these immediate neighbors. Connections made with distant neighbors are, however, penalized by an attenuation factor α. Each path or connection between a pair of nodes is assigned a weight

determined by α and the distance between nodes as α^d

The **Page Rank** is an algorithm used by Google Search to rank web pages in their search engine results. PageRank works by counting the number and quality of links to a page to

determine a rough estimate of how important the website is. The underlying assumption is that more important websites are likely to receive more links from other websites.