RAPPORTI DI RICERCA

Rapporto n. 243

Titolo

Betweenness Centrality and Vertex Relative Position in Subnetworks

Autori

Carlo Drago, Rosanna Grassi

Luglio 2016

Dipartimento di Statistica e Metodi Quantitativi (DiSMeQ)
Università degli Studi di Milano Bicocca
Via Bicocca degli Arcimboldi, 8 - 20126 Milano

Betweenness Centrality and Vertex Relative Position in Subnetworks

Carlo Drago[1], Rosanna Grassi[2]

Abstract

In this work, we address the problem of assessing the node position related to centrality of the other nodes, for both the whole network as well as an identified attribute-based subnetwork. To this end, a methodology is proposed, consisting in computing a relative betweenness measure rather than measuring the node position by means of the betweenness centrality. Our indicator is the ratio of the betweenness value to a power of the total betweenness measure, referred to the subnetwork. A suitable positive parameter α adjusts the intensity of the subgraph betweenness in computing the relative betweenness.

Our methodology make it possible the possibility of characterizing the different nodes based on their position not only as regards the entire network, but also considering their subnetworks identified by a relevant attribute. In order to test our measure we provide examples obtained by the simulation of various network structures. In order to show the effectiveness of our methodology, we propose two possible real applications, in two different fields. The algorithm of this method is provided, in order to allow the replication of the computations.

Keywords:

Betweenness centrality, Subgraphs, Interlocking directorates

Introduction

Complex systems are widely spread in the real world and understanding their functioning is crucial for modeling them in the best way. Network theory is an important tool for describing complex systems and for this reason, it has attracted attention of researchers. It has been used in many fields, such as social, biological, economic and financial contexts (see for instance Wasserman & Faust, 1994, Albert & Barabasi, 2002, Soramaki et al.,

[1] University of Rome "Niccolò Cusano", Italy

[2] Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milano, Italy

Corresponding Author: Carlo Drago, University of Rome "Niccolò Cusano"

2007, Lux & Marchesi, 1999). Many network characteristics are suitable for identifying different features of real complex systems, depending upon which network property needs to be enhanced (Watts & Strogatz, 1998, Fortunato, 2010, Albert & Barabasi, 2002, Newman, 2002, Piepenbrink & Gaur 2013). In real networks, assessing the relative importance of nodes in fundamental in identifying the key elements in the network (for instance in communication and transportation networks, see Agryzkov et al., 2014). Centrality measures are one of the most popular topics in consideration of this issue.

The aim of this paper is to analyse the "relative" centrality of a single vertex with respect to both the whole network as well as a subnetwork, through a relative betweenness centrality measure. Classical centrality measures based on shortest paths quantify the vertex position in the network. Among them, the node betweenness measures the intermediary role of a node based on the frequency in which it belongs to shortest paths existing between any other pairs (Freeman, 1977). In the literature, this (possibly normalized) measure is used for absolute measuring. This can be useful, for instance, when we intend to provide a ranking of nodes according to their position.

However, a node can play an influential intermediary role simply because it belongs to a specific sub-structure of the network (such as a clique or a community), or its role might be not particularly important in its subnetwork, but strategic as concerns the network as a whole. To what extent is the centrality of a node determined by the connections in its subgroup?

In this work, we are not specifically interested in a new ordering of the nodes as regards their individual position, but rather we address the problem of how to assess the node position as it relates to the centrality of the other nodes for both as regards the whole network as well as an attribute-based subnetwork.

With this aim, we propose a methodology consisting in computing a relative betweenness measure rather than measuring the node position by means of the classical measure. Our metric is the ratio of the classical betweenness value to a power of the total betweenness measure, referred to the subnetwork. A suitable positive parameter α adjusts the intensity of the subgraph betweenness in computing the relative betweenness.

Our methodology makes it possible to characterizing the different nodes on the basis of their position not only in relation to the entire network, but also to their subnetwork, which is identified by a relevant attribute.

In a complex system, an attribute is a qualitative node characteristic; quantitatively, it can be described by a categorical or a continuous variable. The different attributes can have a different impact on the structural characteristics, as the centrality. In this perspective, one may reasonably suppose that this has an impact on the betweenness centrality values of nodes aggregated in a subset according to a specific attribute.

Potential applications are possible in several fields, as the problem of quantifying the importance of a node referred to the node position respect to a specific subgraph is frequent in various areas. Considering social networks of co-authors, it could be of interest to classify the importance of some authors solely because they have cooperated on a specific topic although they belong to different departments in various countries.

In interlocking directorship networks, the identification of crucial nodes by means of their financial or non-financial attribute might be essential. If we focus on geographic networks, the importance could be locally related to the region to which nodes belong.

The topic is also of interest in organizational frameworks. Organizations is built upon a complex architecture, and the strength of its ties, as well as how efficiently the information flows within the organization, is fundamental for establishing a company's success.

There are many real examples in which it is important to analyse the structure of subgroups in relation to the whole network, as for example the organizational structure as Departments, hierarchies or organization charts (Brass, 1984, Mehra et al., 2006, Sparrowe et al., 2001). In all these cases, it is important to consider not only the whole network, but also the structure of a network's different parts.

The paper is organized as follows: we described the most used centrality measures (for vertices and subgroups) in Section 2 and the applied methodology in Section 3. In Section 4 some leading examples are presented, in Section 5, two real networks are examined and, in Section 6, we discuss the results and implications for organization studies. Conclusions follows. The proposed algorithm is in the Appendix.

2. Centrality

At first, we briefly recall some standard definitions about graph theory in order to make the reading of the article self-consistent. For detailed definitions see Harary, 1969. A network is a graph $G = (V, E)$, where $V = \{v_1, v_2, \ldots, v_n\}$ is the set of *n* vertices and E is the set of *m* edges (or links). We consider simple, undirected graphs, i.e graph in which edges have no orientation, without self-loops and multiple links. When two vertices share a link, they are called adjacent. The degree d_i of a vertex v_i, $(i = 1, \ldots, n)$ is the number of edges incident with it. A path is a sequence of adjacent vertices $v_1, v_2, \ldots, v_k$ in which all vertices are distinct. The distance $dist(u, v)$ between two vertices u and v is the length (intended as the number of links) of the shortest path from u to v ($u-v$ geodesic). A graph is connected if for each pair of vertices u and v there is a path connecting u and v, otherwise is disconnected. In this case, it has more than one connected component.

Centrality is one of the most studied concepts in network analysis. Vertex centrality measures the importance of a vertex's position in a network. Various measures of centrality exist, depending on their different interpretations and their different purposes of application. A node could be highly central if it is adjacent to many other nodes, or if it is important as an intermediary vertex in communication among others.

The idea of centrality has been also extended to other network elements. For instance, the betweenness of an edge is the fraction of shortest paths between pairs of vertices that run along it. Edge betweenness is used in the well-known Girvan-Newman algorithm in order to detect communities in a network (Girvan & Newman 2002).

A possible extension of the concept of vertex centrality is the centrality of a subset of vertices, in particular subnetwork aggregated by certain properties.

In this section, we recall the most used vertex centrality measures and the centrality of subset of vertices (subgraphs). Special attention is paid to betweenness centrality measure, which is the measure we have used to perform our analysis.

Vertex centrality measures

Different indicators have been used to quantify the centrality, depending on the topological features of the network. In the last few years, the literature on network centrality has continued to be enriched by new results. In this section, we just mention the most known measures; nevertheless, this is not an exhaustive list.

Two classical measures are based on the adjacency relation of a node. The first one is the degree centrality d_v of the node $v \in V$; this is the number of the neighbours of the node, then it is the most immediate and intuitive measure of node centrality.

The second one is the eigenvector centrality x_i (see Bonacich, 1972) :

$$x_i = \frac{1}{\rho} \sum_{i=1}^{n} a_{ij} x_j,$$

where a_{ij} is the $(ij)^{th}$ entry of the adjacency matrix **A** and ρ is its spectral radius. It is well known in matrix theory (see Horn, 1985) that all components of **x** are real and strictly positive. This measure is the sum of its direct connections weighted by their own centralities, and acts as a measure of the influence and power of a node with respect to the overall network. The normalized eigenvector score is $\frac{x_v}{\|x\|_2}$, where $\|.\|_2$ is the Euclidean norm[3].

Vertex betweenness centrality, defined in Freeman, 1977, counts the fraction of shortest paths between a pair of nodes that an intermediate node v lies, summing these fractions over all node pairs:

$$b(v) = \sum_{i<j} \frac{g_{ij}(v)}{g_{ij}}, \qquad i, j \neq v$$

where g_{ij} is the number of $i-j$ geodesics, and $g_{ij}(v)$ is the number of $i-j$ geodesics passing through v.

This measure can be usually normalized to lie between zero and one, by dividing the betweenness value $b(v)$ by its maximum value: $b_N(v) = \frac{b(v)}{\binom{n-1}{2}}$.

Betweenness is suitable for measuring the importance of a vertex in terms of its intermediary position in the network when the flow travels along geodesics. This measure is related to the frequency of paths passing through the vertex. However, in real networks information frequently does not flow only along geodesic paths.

Several other measures have been proposed in the literature, that take all paths into account, and not only the shortest ones, in order to provide a more realistic betweenness measure. In some sense, these measures are founded on the many different ways to connect a pair of nodes: the most important ones are respectively flow and random walk betweenness.

Flow betweenness (Freeman, Borgatti & White, 1991) extends the betweenness idea to weighted graphs, including the contribution of non-geodesic paths, and it is related to the concept of maximum flow, computed by using the Ford-Fulkerson algorithm (Ford & Fulkerson 1956):

$$fb(v) = \sum_{i<j} m_{ij}(v) \qquad i, j \neq v$$

where $m_{ij}(v)$ is the maximum flow from node k to node j passing through i. The normalized flow betweenness is obtained by dividing the flow betweenness value $fb(v)$ by the total flow through all pairs of vertices different from v; this normalized measure ranges

[3] The Euclidean norm is defined as $\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$.

between 0 and 1 and gives the proportion of the flow that depends on i. Thus, the flow betweenness measures the contribution of a node to all possible maximum flows.

Random walk betweenness was proposed in Newman, 2005 and it is based on random walks, counting how often a node is traversed by a random walk between two other nodes. This measure is, in some sense, the opposite of that introduced by Freeman: betweenness is suitable for situations in which information travels along a precise route (the shortest one), whereas the random betweenness is suitable for a network in which information travels randomly until it finds its target. Moreover, it includes contributions from many paths that are not optimal in any sense, although shorter paths still tend to count for more than longer ones since it is unlikely that a random walk becomes very long without finding the target.

In introducing the measure, Newman 2005 first uses the analogy of the flow of electrical current in a circuit, then he show that this is also equivalent to the flow of a random walk. Random walk betweenness is still extensively used and studied.

Several measures aim to establish the central vertex position in terms of "how close" the node is to others. These centralities are then directly related to the idea of proximity of the node to others.

The closeness formula is based on the reciprocal of the sum of the distance between a vertex and all others:

$$c(i) = \frac{1}{\sum_{\substack{j \in V \\ j \neq i}} d(i,j)}$$

where $d(i,j)$ is the geodesic distance between nodes i and j. This measure can be normalized, by multiplying the score by $n-1$.

From this point of view, agents that are reachable by all others through shortest paths are more important, due to the fact the effort to transmit information using geodesic patterns is little. A similar measure, called efficiency, has been proposed by Latora and Marchiori in (Latora & Marchiori, 2007) in order to measure the mean flow-rate of information over the network.

Centrality measures referred to subgraphs

A natural extension of the concept of vertex centrality is the centrality of a subset of vertices and several definitions of group centrality have been introduced in the literature.

In Everett & Borgatti, 1999, the concept of group centrality is defined as an extension of the existing measures for individuals (degree, closeness and betweenness centrality), making it possible to identify what groups or sets of nodes are more prominent in a network, but also, given a network of ties among organization members, how a team that is maximally central can be formed. In particular, the betweenness of a set is defined in terms of shortest paths passing thorough at least one vertex of the set (group betweenness centrality).

Grassi et al., 2008 define a measure of set betweenness as the sum of the betweenness values of the vertices in the subset (total betweenness). Kolaczyk et al. 2009 propose a measure of set betweenness in terms of shortest paths passing through all vertices of the set (co-betweenness centrality). Other works extend the centrality to particular subgroups in the network, as core/periphery structures or two-mode graphs (see Everett & Borgatti 2005).

Some works are based on an idea of relative centrality, i.e. a vertex centrality, related in some way to a set or a group of nodes.

Estrada and Rodríguez-Velázquez 2005, define a vertex centrality, according to the number of weighted closed walks starting and ending at the node. Given thatevery closed walk is weighted their influence on the centrality decreases as the order of the walk increases. Each closed walk is associated with a connected subgraph, which means that this measure counts the times that a node takes part in the different connected subgraphs of the network, with smaller subgraphs having higher importance. Consequently, they call this measure the subgraph centrality for nodes in a network. In Smyth & White 2003, the authors propose a measure of node importance in a graph with respect to a set of nodes (Markov centrality). They also propose a general class of algorithms for relative importance based on weighted paths and motivated by graph-theoretic ideas and they discuss experimental results on three real-world networks and correlation of ranks from different algorithms is analysed. Wang et al. 2013 address the problem to measure the relative importance between two nodes and they present a measure (path probability) to represent the connection strength between the ending and a starting node. The measure proposed use the sum of the path probabilities of all the important paths between the two nodes. Bell 2014, deals with the problem of defining a measure, called subgroup centrality, of both local and global influence of a node relative to some subset of nodes.

3. Methodology

The idea of defining a measure that refers to centrality of a subnetwork, instead of a single node, is useful when we are interested in establishing how important a vertex is in as much as apart of an eminent group of vertices into the network. Some vertices might not be important *per se* but rather because their role into a significant group or coalition. In this regard, it might happen that a node is not important in the entire network, but its role completely changes if it is compared to a subset of nodes characterized by one, or more, aggregate attributes.

Both measures - global and local - are important to be simultaneously investigated because the role of a node could change significantly depending not only on the quantity and quality of its direct links (to be more precise, depending on how many and who its neighbours are) but also on the part of the network in which this node is located.

To analyse the relative position on a node with respect to centrality of a subset of vertices to which the node belongs, we use the following measure:

$$b_R(v) = \frac{b(v)}{(b^T)^\alpha},$$

where $b(v)$ is the betweenness of the individual vertex v and b_T is the total betweenness, referred to the subset. The measure we propose is a relative measure of the intermediary role of the vertex in the network, which takes into account in which subnetwork this vertex is located.

In order to enhance the importance of the vertex relative position into the subnetwork we use a suitable real parameter $\alpha \in [0,1]$. The a-power of the total betweenness quantify the importance of the betweenness of the subgraph in computing the relative betweenness. Indeed, if $\alpha = 0$, the proposed measure is equal to the classical vertex betweenness, whereas if $\alpha = 1$ we take into account of the intermediary role of this vertex embedded into the subnetwork. Values of α between 0 and 1 attenuate the "subgroup effect" in vertex centrality, hence, α can be suitably tuned in computing the betweenness of a vertex, giving more (or less) importance also to the total betweenness of the subgroup to which the vertex belongs. Practical uses of this measure imply that it is necessary to specify the value of the parameter α, in accordance, for instance, with the relevance of the subgroup attribute. In this perspective, the question concerning the best value of α for a given network should be investigated. In this article we do not address this problem, and both simulation and empirical results are analysed assuming $\alpha = 1$, (the case of maximum "subgroup effect").

Special attention have to be paid when the total betweenness vanishes. This happens in extremal cases, when one extracts a complete subgraph (or with more than one component, each a complete subgraph) or composed by isolated nodes. In these specific situations, our measure cannot be applied. However, in these specific cases, uncommon in real networks, also the betweenness score of a single vertex is equal to zero.

It is worth noting that by means of our methodology we are not only proposing a new ranking between the vertices, a topic that has been already taken in the literature by means of different from centrality measures.

Our aim is different. We are proposing a new measure for evaluating sophisticated characteristics of the nodes that are difficult to quantify. A measure of a prominence, or centrality, with respect to a connected group of nodes, that could be different from to the centrality compared to the entire network.

In the real world, the single connected group might prove tobe more important than the entire network so, it is possible to expect that centrality might be more relevant or different for single groups than in the entire network as well. Betweenness centrality usually is computed by referring to a specific element of the network, for instance a vertex. However, a vertex position also depends on its position with respect to the subnetwork in which the vertex is located. When we compute classical vertex betweenness we focus only on the individual score, not considering the contribution of the interconnections between the nodes and the group it belongs.

Our analysis is based on the characteristics of the nodes and in particular on their attributes. However, other network characteristics (communities, for instance) can be related to some topological structures of the networks. In our case, the results obtained are different from those obtained by considering the communities extracted by the network.

In the next subsection, we show some simple leading examples to explain the relevance of our methodology.

4. Leading examples

In order to show some examples of our approach, we have considered three networks obtained by simulation. Networks have been designed in order to create three leading examples which are useful to demonstrating the methodology in practice. Networks were generated using the Igraph package of the programming R language of (see Csardi and Nepusz, 2006). We have analysed networks based on three distinct models:

1. Barabasi Model (Barabasi Albert, 1999) in Figure 1 (a)
2. Erdős Renyi Model (Erdős Renyi, 1959) in Figure 2 (a)
3. Forest Fire Model (Leskovec, Kleinberg, Faloutsos 2007) in Figure 3 (a)

The Barabasi Albert model allows us to simulate random scale-free networks, whereas the Erdős-Renyi model generates random graphs networks. Forest Fire Model makes it possible to simulate a scheme on real networks in which the attachment mechanism occurs at the network periphery and not at its core. Hence, in this sense the Core-Periphery structure seems to be very relevant to this case.

The networks have 30 nodes and are characterized by nodes with one out of three different attributes (which are indicated as "1" "2" and "3"). In each network, we have determined the different subnetworks by identifying the attributes of the connected nodes. For example, for the first network (Figure 1 (a)), we have identified three specific subgraphs (Figure 1 (b), (c) and (d)), on the basis on the attributes given. In these examples, we consider an attribute for each node, namely a categorical variable with three different values.

The algorithmic procedure consists at first in extracting the attribute-based subnetwork, then in computing the relative centrality $b_R(v)$. The subnetwork is considered in this work as a network separated by the whole structure and the relative betweenness is computed in the algorithm accordingly.

Some relevant differences can be detected on the simulated networks. In the first case (the Barabasi model) the network appears to not be particularly dense, unlike the second case (the Erdős-Renyi model), whereas the third case represents a mix of the other two networks the Forest Fire Model).

Few nodes, located in a central position, characterize the first network so we then expect a strong advantage from the most central nodes. In the second network we can find various different vertices showing high betweenness scores. The third network shares some relevant characteristics with the first and the second network. In this case as well the network presents different nodes having a high betweenness, located at the "core", whereas the nodes in the periphery have low betweenness values.

The subgraphs are extracted by taking into account the different attributes. The three subgraphs of the three networks are structurally different from each other, especially the first and the second one. The subgraphs extracted from the third network (Figure 3) exhibit characteristics that seem to be a mix between the ones extracted by the other two networks (Figure 1 and 2).

It is important to note that the different subgraphs, obtained by considering the attributes of the nodes of the global network are not simply identifiable by inspection of the entire network. Indeed, the visualization of the subgraphs characterized by node attributes is usually hidden in the entire network.

In this sense, the attribute can represent a privileged point of view for observing some interesting patterns, which can be detected in the data, especially where it is not simple to be recognized them in a different way.

The relative betweenness values[4] show some interesting new features of the subgraphs obtained by considering the attribute. In particular, the nodes reflect, through their relative centrality related to a specific attribute, some subgraph characteristics, which are not evident in the whole networks. It is possible to note, in this context, that the position of the subgraph is also relevant. When the subgraph is, in someway, central, it could be identified with the centre of the network, or it can be identified also with the *core*. The centrality of the subgraph is determined by the related attribute (which in this sense is dominant). Relative betweenness (considered on the basis of its ranking) can assess this result, highlighting the most dominant nodes.

On the contrary, the subgraph centrality can be induced by latent variables that are also relevant. In this sense, our methodology can be used to analyse the impact of these latent variables on the network. This fact could prove very useful in organizational networks because it allows us to grasp the hierarchical patterns. Various examples can be considered: first of all, it is possible to detect the case of cooptation[5] on a network of a set of nodes that are specifically linked inside the subnetwork. As an example, in alumni-network, being part of the same set can give rise to the opportunity, for some nodes, to share similar important positions in firms. In this case the latent attribute can be "being part" of an alumni network, for example of a prestigious school. A similar mechanism occurs in the subgroup identified by family firms. In this case, the latent variable is being part of the family and the family members tend to have a position in the firm, becoming nodes of the subnetwork.

In order to analyse the different results for the networks and the subgraphs, we compare the results for the betweenness and the subgraphs at the same time.

[4] The relative betweenness has been computed separately for each connected component of the network.

[5] To be admitted into a group (as a board) by the votes of the existing members.

Some interesting features of the networks can be revealed by the analysis of the most central nodes on the network respect to both the betweenness and relative betweenness. In addition, there are relevant differences between the centrality for the subnetworks and the network, respectively.

In the Barabasi network there are two nodes highly central (nodes 2 and 6, with different attributes) which are also highly central on their respective subgroups. On the contrary, the node 1 is highly central in the network but not in its subgraph. Therefore, there is not a unique subnetwork of the most central nodes, but the nodes are respectively the most central on their groups (Table 1 and 2).

As we have already notice, in the case of the ER network (table 3 and 4) and the FF network (table 5 and 6) we observe a different structure of the network centrality, in particular that referred to the betweenness values.

In the ER network, the three subgraphs reveal a similar homogeneous structure (Figure 2b, 2c, 2d). On the contrary, the FF network can be well divided in high-centralized structures. This is evident from Figures 3b, 3c and 3d in which we are able to immediately identify these central nodes. In the ER network, nodes 1, 23 and 26 are highly central in both network structures. The nodes 10 and 3 are more central in the entire network than in their subnetworks. On the contrary, the nodes 17 and 30 are more central than in the whole network, then they play a strategic role in their respective subnetwork (Tables 3 and 4). A similar analysis can be done for the FF model. Nodes 15 and 25 are highly central, maintaining their position in their respective subnetworks, whereas nodes 2 and 5 achieve importance simply being part of the same subnetwork.

In these cases, there is no a unique subnetwork overlapping the “centre” of the network, consequently, it is possible to deduce that the considered attributes are not dominant in the whole network.

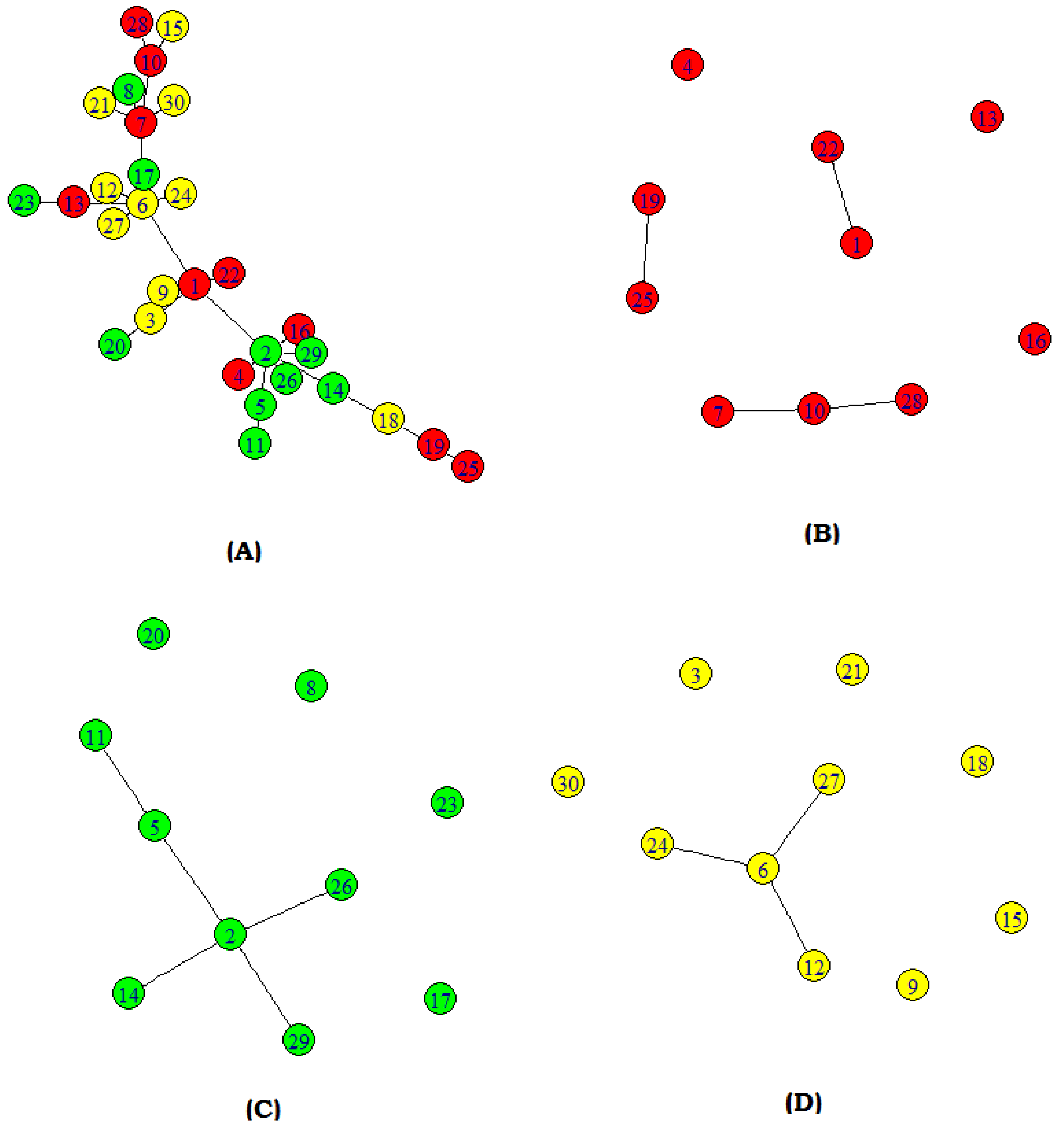

Figure 1: Simulated network using Barabasi Model. The node colours differ accordingly with three different attributes.

Id-Subnetwork 1	Rel Betw	Id- Subnetwork 2	Rel Betw	Id – Subnetwork 3	Rel Betw
1	0.000	2	0.692	3	0.000
4	0.000	5	0.308	6	1.000
7	0.000	8	0.000	9	0.000
10	1.000	11	0.000	12	0.000
13	0.000	14	0.000	15	0.000
16	0.000	17	0.000	18	0.000
19	0.000	20	0.000	21	0.000
22	0.000	23	0.000	24	0.000
25	0.000	26	0.000	27	0.000
28	0.000	29	0.000	30	0.000

Table 1: Relative betweenness scores (subnetworks extracted by BA model)

.

Id	Attribute	Betweenness	Id	Attribute	Betweenness	Id	Attribute	Betweenness
1	1	0.981	11	2	0.000	21	3	0.000
2	2	0.864	12	3	0.000	22	1	0.000
3	3	0.106	13	1	0.106	23	2	0.000
4	1	0.000	14	2	0.295	24	3	0.000
5	2	0.106	15	3	0.000	25	1	0.000
6	3	1.000	16	1	0.000	26	2	0.000
7	1	0.568	17	2	0.000	27	3	0.000
8	2	0.000	18	3	0.205	28	1	0.000
9	3	0.000	19	1	0.106	29	2	0.000
10	1	0.208	20	2	0.000	30	3	0.000

Table 2: BA network: Id, attributed and betweenness

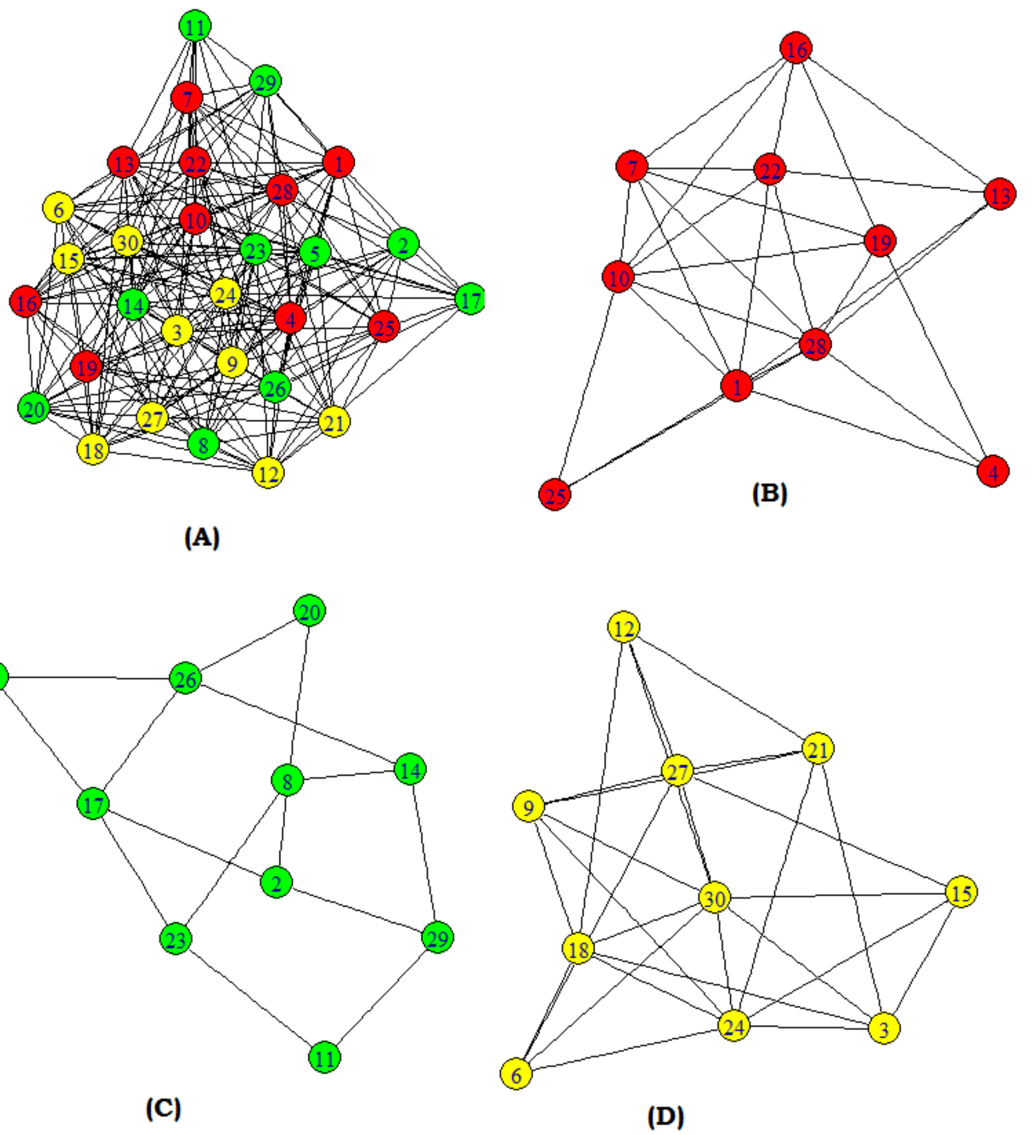

Figure 2: Simulated network using Erdős-Reny Model. The node colours differ accordingly with three different attributes.

Id- Subnetwork 1	Rel. Betw.	Id- Subnetwork 2	Rel. Betw	Id – Subnetwork 3	Rel. Betw.
1	0.213	2	0.086	3	0.062
4	0.014	5	0	6	0.025
7	0.053	8	0.171	9	0.050
10	0.173	11	0.028	12	0.039
13	0.025	14	0.109	15	0.030
16	0.069	17	0.194	18	0.116
19	0.104	20	0.021	21	0.094
22	0.053	23	0.146	24	0.173
25	0	26	0.162	27	0.169
28	0.296	29	0.083	30	0.243

Table 3: Relative betweenness scores (subnetworks extracted by ER model)

Id	Attribute	Betweennes s	Id	Attribute	Betweennes s	Id	Attribute	Betweennes s
1	1	0.587	11	2	0.168	21	3	0.466
2	2	0.369	12	3	0.592	22	1	0.712
3	3	0.801	13	1	0.627	23	2	0.819
4	1	0.765	14	2	0.787	24	3	0.075
5	2	0.459	15	3	0.464	25	1	0.515
6	3	0.392	16	1	0.488	26	2	0.662
7	1	0.464	17	2	0.217	27	3	0.305
8	2	0.420	18	3	0.343	28	1	0.674
9	3	0.692	19	1	0.321	29	2	0.364
10	1	1.000	20	2	0.383	30	3	0.657

Table 4: ER network: Id, attributed and betweenness

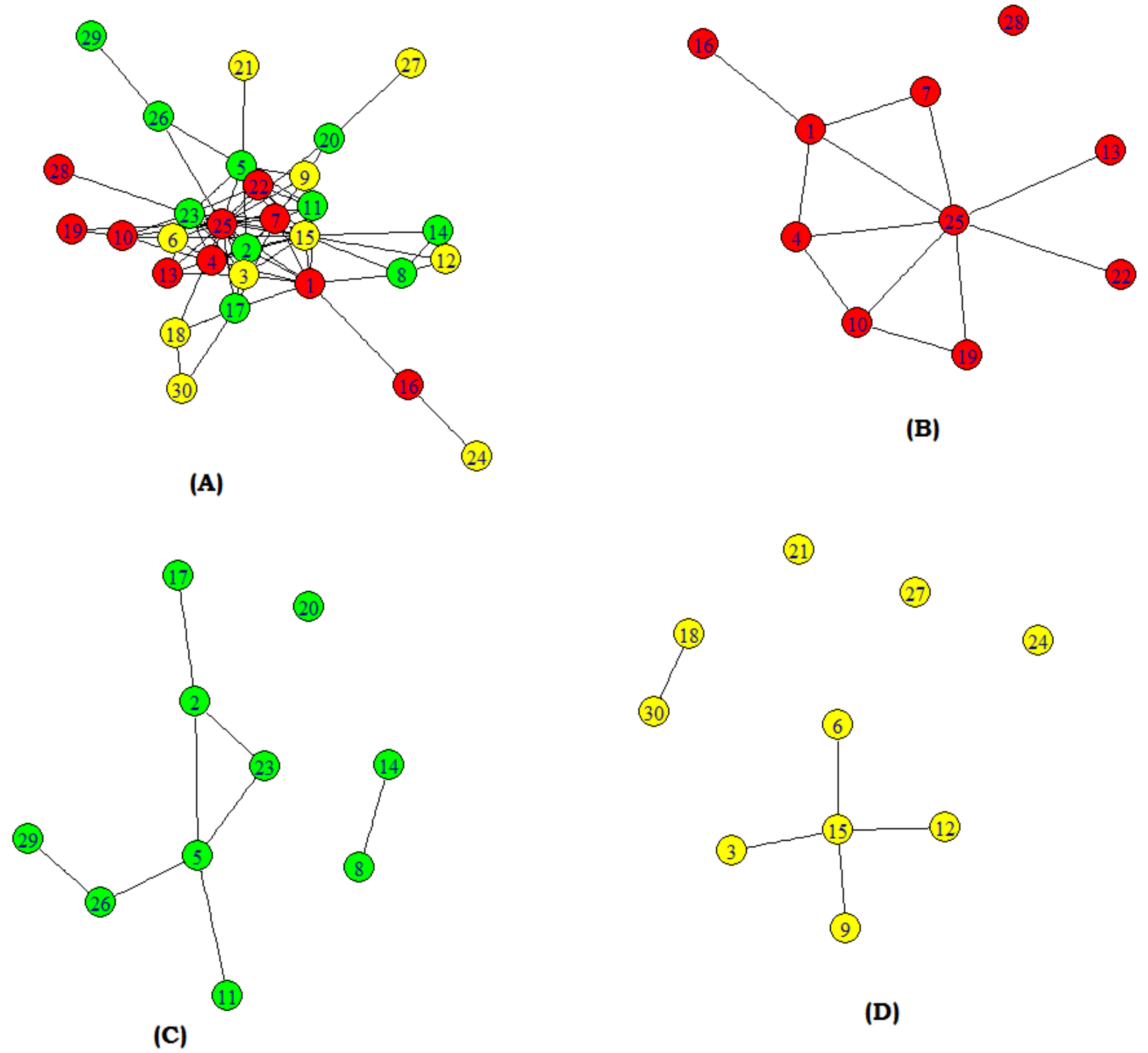

Figure 3: Simulated network using Fire Forest Model. The node colours differ accordingly with three different attributes.

Id- Subnetwork 1	Rel. Betw.	Id- Subnetwork 2	Rel. Betw	Id – Subnetwork 3	Rel. Betw.
1	0.268	2	0.238	3	0
4	0.036	5	0.524	6	0
7	0	8	0	9	0
10	0.018	11	0	12	0
13	0	14	0	15	1
16	0	17	0	18	0
19	0	20	0	21	0
22	0	23	0	24	0
25	0.679	26	0.238	27	0
28	0	29	0	30	0

Table 5:: Relative betweenness scores (subnetworks extracted by FF model)

Id	Attribute	Betweenness	Id	Attribute	Betweenness	Id	Attribute	Betweenness
1	1	0.433	11	2	0.026	21	3	0.000
2	2	0.079	12	3	0.000	22	1	0.003
3	3	0.021	13	1	0.000	23	2	0.258
4	1	0.037	14	2	0.000	24	3	0.000
5	2	0.250	15	3	0.493	25	1	1.000
6	3	0.006	16	1	0.179	26	2	0.179
7	1	0.160	17	2	0.146	27	3	0.000
8	2	0.024	18	3	0.046	28	1	0.000
9	3	0.000	19	1	0.000	29	2	0.000
10	1	0.013	20	2	0.179	30	3	0.000

Table 6: FF network: Id, attributed and betweenness

5. Real network application

Bank Wiring Room network

In order to evaluate our methodology in a real framework, we have considered two real networks. The first one describes the Bank Wiring Room data of Roethlisberger and Dickson (1939), available in UCINET dataset (Borgatti, Everett & Freeman, 2002), represented in Figure 4.

We have excluded the isolated nodes (I3 and S2). The size of the nodes is visualized in different manners, according to their betweenness values.

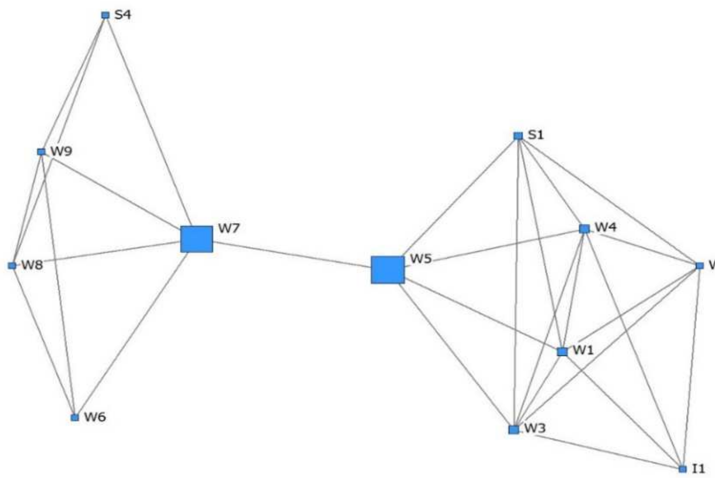

Figure 4: Graph of Bank Wiring Room data.

The betweenness values of the vertices suggest that vertices W7 and W5 play a highly central role. The nodes I1,S1,W1,W2,W3,W4 and W5 form a cohesive group, where the node W5 is clearly the most important. Let us call G' the subgraph to which these nodes belong. Quite surprisingly, the score of I1 and W2 is low with respect to the one of W1,W3 and W4. The total betweenness value (with respect to the entire network) is $b_T = 0{,}923$. The relative betweenness is computed for $\alpha = 1$ in order to fully enhance the effect of the node position in the subnetwork. Computing the relative betweenness, the ranking is completely confirmed (see Table 7).

Vertex	Betw	Rel Betw
W5	0.384615385	0.4166725
W1	0.048076923	0.0520841
W3	0.048076923	0.0520841
W4	0.048076923	0.0520841
W7	0.36324359	0.3935194
S1	0.019230769	0.0208336
W8	0.004269231	0.0046251
W9	0.004269231	0.0046251
W2	0.003205128	0.0034723
I1	0	0
S4	0	0
W6	0	0

Table 7: Graph of Bank Wiring Room data: betweenness and relative betweenness with respect to the whole network.

We now compute the relative betweenness with respect to the subgraph G'. To this end, we first extract from G the subgraph G', then we calculate the vertex betweenness values. The total betweenness value, referred to the subgraph G', is $b(G')_T = 0{,}19993$. Table 8 provides the values of betweenness and relative betweenness in this case.

Looking at the position of same nodes with respect to the subgraph G', the ranking obtained is different. The position of nodes W1,W3 and W4 has totally changed becoming now the highest central, whereas the score of W5 vanishes. The relative betweenness also in this case preserves the ranking among nodes. However, their value has been reduced in regard of the classical betweenness and the position of the nodes W1,W3 and W4 is now scaled down, due to the effect of the presence of many nodes with highest value (given that the parameter α is equal to 1, this effect is totally evident).

Vertex	Betw	Rel Betw
W1	0.833	0.277759
W3	0.833	0.277759
W4	0.833	0.277759
S1	0.25	0.083361
W2	0.25	0.083361
I1	0	0

W5	0	0

Table 8: Subgraph of Bank Wiring Room data: betweenness and relative betweenness referred to the subgraph G'.

Interlocking directorship network in Italy

As a second application of our method, we consider the interlocking directorate network in the case of Italy referred to the year 2006.

In the past, there was a growing volume of network research in Management (Borgatti and Foster 2003). In this context, the interlocking directorate networks are an important issue for their implications for the information flow among the different firms (Nooteboom 1999). Interlocking directorate networks grows for many different reasons (Mizruchi 1969 Schoorman, Bazerman, Atkin 1981) and they exists as communication channel between the different firms (Gulati 1998). The interlocking directorates lend themselves to being an instrument both for cooperation as well as forcommunication between the different companies (Palmer 1983, Koenig et al. 1979) and they can be used at the same time for monitoring purposes (Dooley 1969). In this sense, the financial companies typically tend to monitor the non-financial companies as well (for example of interlocking directorates for the purpose of monitoring by banks see Mariolis 1975). A third reason for an interlocking is collusion (Hansen 2014 Santella et al. 2006).

The monitoring theoretical explanation arises from the fact that some specific companies need to monitor other companies (for instance the banks). In this sense, the monitoring can be exercised by means of the interlocking directorate

Similarly, also the cooperation can be obtained in different ways, for example in the case of firms which are interconnected because of their buyer- seller relationship. In these cases, it could be relevant to share one, or more directors. The interlocking directorate network shows some important characteristics (Piepenbrink & Gaur 2013), as small-worldness (see for instance Robins & Alexander, 2004) or being structured in different modules or communities, usually interrelated to some advanced forms of cooperation.

In these kinds of network, it is important to understand the structure in relation to the firm's different characteristics. For instance, financial companies tend to monitor their credits towards others financial firms and to connect themselves to non-financial firms. They usually play a prominent role in the interlocking directorate networks and it might consequently be important to analyse their structure in the network.

For these reasons, have extracted the relevant information of the subnetwork of the financial companies for this structure. The subnetwork can show some of the competitive behaviours of the connected companies.

As we have already noted in the previous section, the position of the subgraph in the network is relevant. A significant attribute can make central the subgraph that can correspond to the core of the network. In this situation, the absolute betweenness (referred to the entire network) and the relative betweenness tend to share the same values.

We consider the network based on the specific connections between the different companies (in this case, the Italian companies that are highly capitalized on the market). The nodes of the network are the companies and the links connecting the companies are the directors. The data used are the same as in Santella et. al 2006 and they are related to the SP MIB companies, which are the most capitalized companies of the Italian market. The network generated is depicted in Figure 5.

The Italian case shows some peculiar characteristics. We have considered the firm's typology (financial and non-financial) as an attribute. This attribute is very important because allows to study both the mechanisms of competition as well as collusion among the various companies.

We have extracted, the subnetwork of the financial companies related to the SP MIB (Figure 7) and the non-financial companies (Figure 6) from the entire network and we have computed the relative betweenness.

It is essential to compare the results obtained from both the subnetwork and the global network. The results show the different structure of the subnetwork as regard the whole network. This does not occur when the subnetwork overlaps the core of the network, then we can conclude that this is not our case

However, there are some companies, which show a similar result for both the networks. We have observed that the financial companies display higher centrality values in the subnetwork than in the global network structure.. Concerning this matter, it should be notedthat financial companies play a crucial role throughout the entire network of the Italian corporations, therefore it is important to analyse the subnetwork results in order to understand Italian capitalism. At the same time, it is possible to note that some companies perform well in both the financial subnetwork as well as in the whole network, thus confirming their high centralities in the structure.

It should be notedthat this analysis describes a very specific situation (concerning the financial sector) in a given time. After the reform[6] occurred in Italy in the 2011, the financial

[6] The interlocking directorate reform of 2011 (included in the so-called "Save Italy" Decree) outlaws the interlock among financial firms. A director of a financial company, who served on the board of two (or more) financial firms, had to exercise the choice (option) between one of the two (or more) directorates by April 27, 2012 otherwise, he would lose the appointments.

companies subnetwork of the disappears. The attention paid by the Italian government to this matter confirms the relevance of the dominance as a concept. This method is therefore particularly important for identifying the various relevant subgraphs and their dominant nodes.

We can now compare the relative betweenness node centrality in the specific subnetwork to the node centrality in the whole network. Various situations are possible: the case of a single node which is prominent in the entire network (high relative betweenness considering the entire network), but not in the subnetwork, as well as the case of a node which is prominent (high relative betweenness) in the subnetwork but not in the network.

The relative betweenness in this context it is particularly useful because allows us to identify nodes which are in a particular position inside their subnetwork. There are cases in which the attitude to be central on the network is specifically related to a defined attribute. In this case, the centrality is determined by a specific attribute and the other attributes do not have a strong impact on the centrality of the single node.

In particular, the prominent role of Mediobanca emerges among the centrality of the financial companies, which is characterized by a very high relative betweenness, but one that is not so high in consideration of the entire network. On the contrary, is the case of Unicredit. This bank displays a high relative centrality in the network and a low centrality in the financial network. That might be explained with the fact that Unicredit shows many connections with non-financial companies and its centrality consequently increases in the global network.

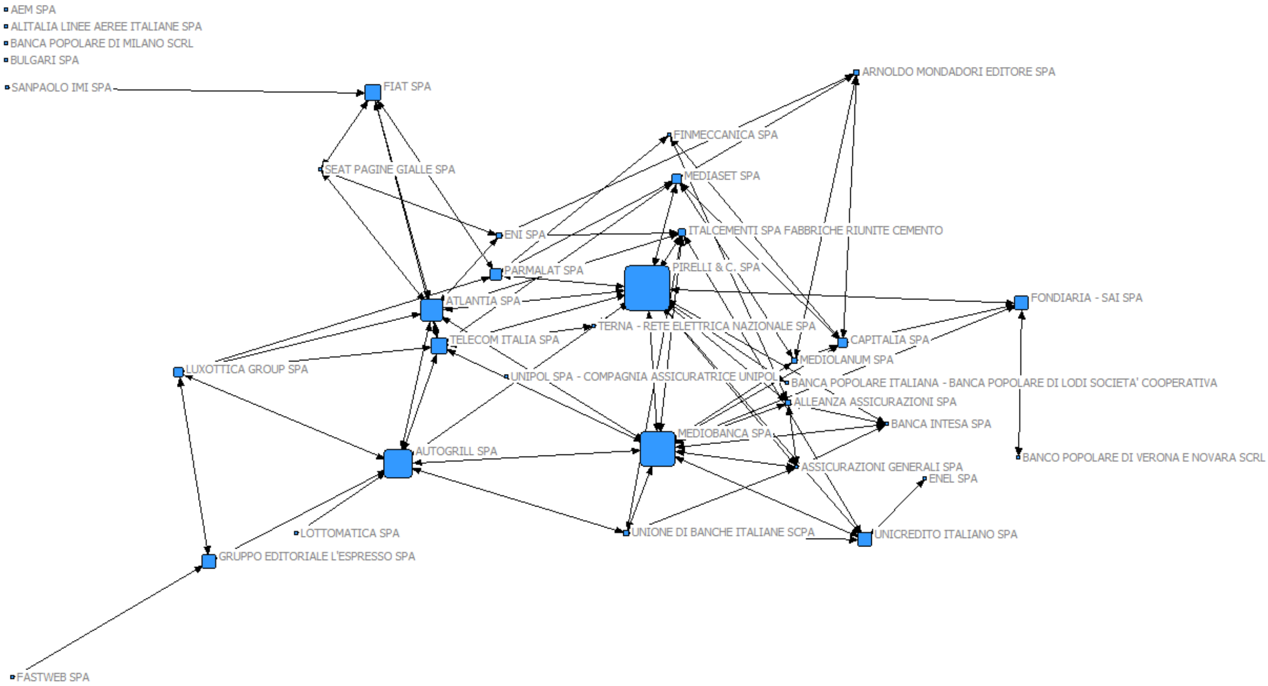

Figure 5: Interlocking directorship network SPMIB 2006

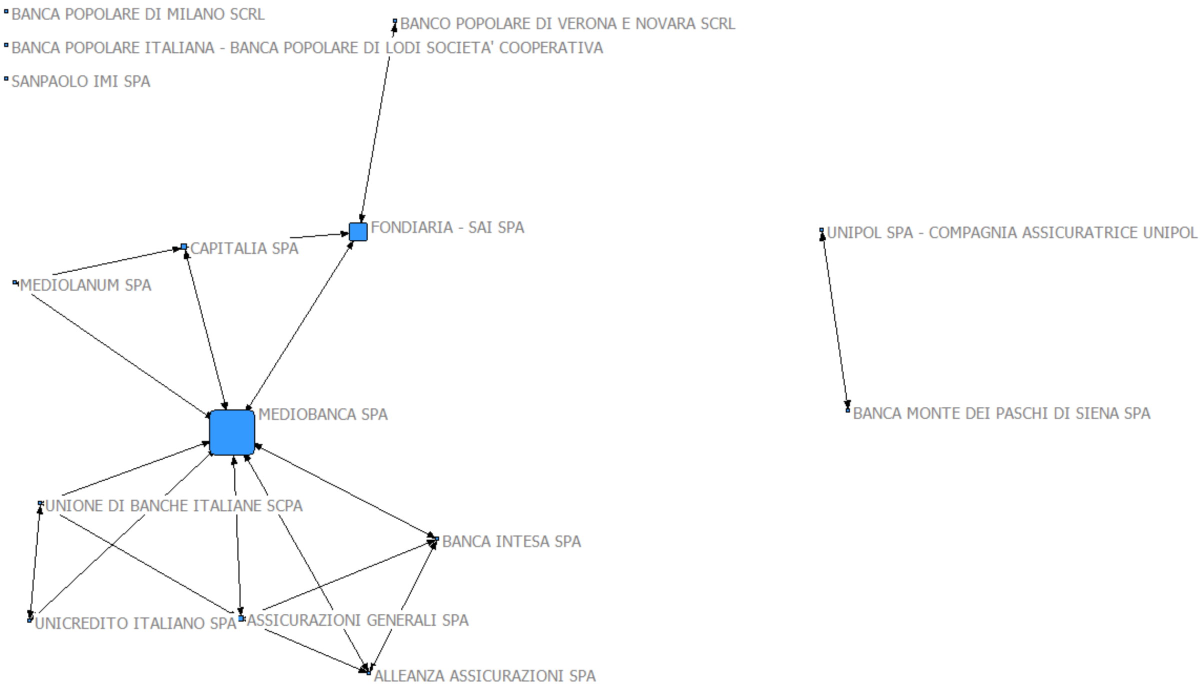

Figure 6: Subgraph of Financial Companies

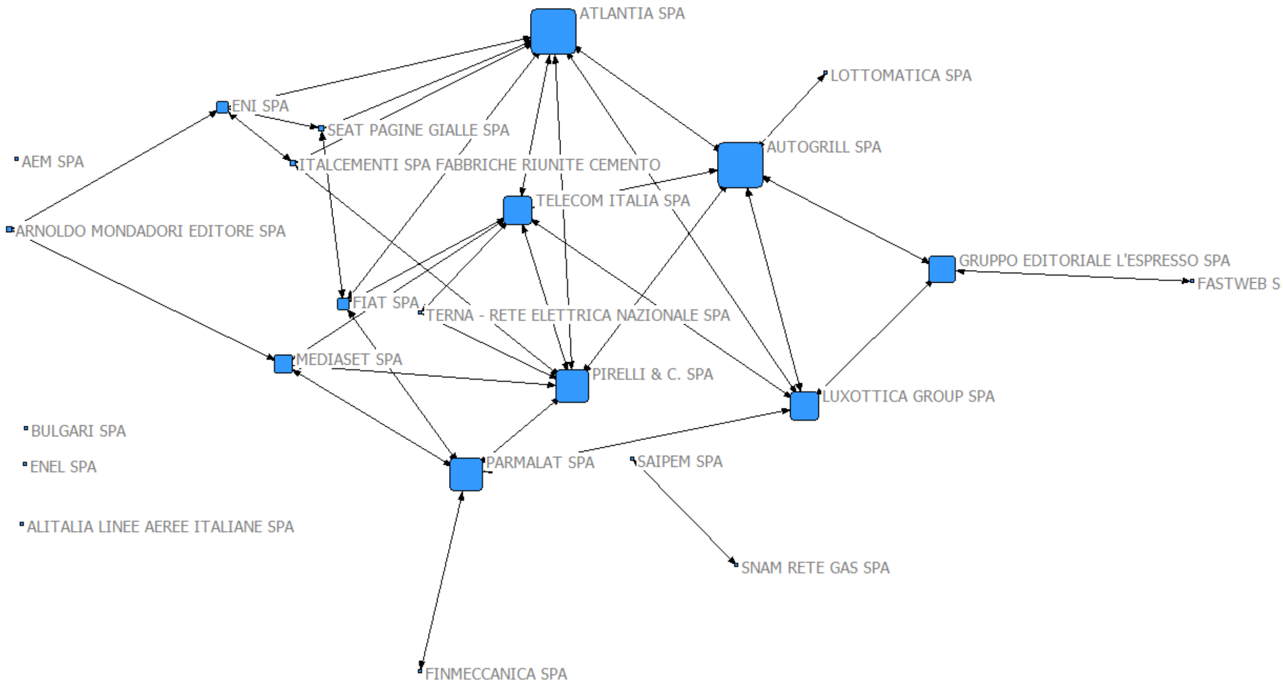

Figure 7: Subgraph of Non-financial Companies

Id	Degree	Betw.	Clos.	Financial	Relative betw.
AUTOGRILL SPA	6	1.000	189	0	0.165
ATLANTIA SPA	8	0.981	187	0	0.162
PIRELLI & C. SPA	7	0.725	187	0	0.12
PARMALAT SPA	5	0.709	191	0	0.117
LUXOTTICA GROUP SPA	5	0.600	189	0	0.099
TELECOM ITALIA SPA	7	0.575	188	0	0.095
GRUPPO EDITORIALE L'ESPRESSO SPA	3	0.561	199	0	0.093
MEDIASET SPA	4	0.362	194	0	0.06
ENI SPA	4	0.220	197	0	0.036
FIAT SPA	4	0.190	194	0	0.031
ARNOLDO MONDADORI EDITORE SPA	2	0.057	203	0	0.009
ITALCEMENTI SPA FABBRICHE RIUNITE CEMENTO	3	0.044	195	0	0.007
SEAT PAGINE GIALLE SPA	3	0.041	197	0	0.007
AEM SPA	0	0.000	529	0	0
ALITALIA LINEE AEREE ITALIANE SPA	0	0.000	529	0	0
BULGARI SPA	0	0.000	529	0	0
ENEL SPA	0	0.000	529	0	0
FASTWEB SPA	1	0.000	214	0	0
FINMECCANICA SPA	1	0.000	206	0	0
LOTTOMATICA SPA	1	0.000	204	0	0
SAIPEM SPA	1	0.000	507	0	0

SNAM RETE GAS SPA	1	0.000	507	0	0
TERNA - RETE ELETTRICA NAZIONALE SPA	2	0.000	199	0	0
		6.064			

Table 9: Centrality measures and relative betweenness scores (non-financial companies)

Id	**Degree**	**Betw.**	**Clos.**	**Financial**	**Relative betw.**
MEDIOBANCA SPA	8	1.000	100	1	0.7
FONDIARIA - SAI SPA	3	0.327	105	1	0.228571
ASSICURAZIONI GENERALI SPA	4	0.041	105	1	0.028571
CAPITALIA SPA	3	0.041	105	1	0.028571
UNIONE DI BANCHE ITALIANE SCPA	3	0.020	106	1	0.014286
ALLEANZA ASSICURAZIONI SPA	3	0.000	106	1	0
BANCA INTESA SPA	3	0.000	106	1	0
BANCA MONTE DEI PASCHI DI SIENA SPA	1	0.000	211	1	0
BANCA POPOLARE DI MILANO SCRL	0	0.000	225	1	0
BANCA POPOLARE ITALIANA - BANCA POPOLARE DI LODI SOCIETA' COOPERATIVA	0	0.000	225	1	0
BANCO POPOLARE DI VERONA E NOVARA SCRL	1	0.000	113	1	0
MEDIOLANUM SPA	2	0.000	107	1	0
SANPAOLO IMI SPA	0	0.000	225	1	0
UNICREDITO ITALIANO SPA	2	0.000	107	1	0
UNIPOL SPA - COMPAGNIA ASSICURATRICE UNIPOL	1	0.000	211	1	0
		1.429			

Table 10: Centrality measures and relative betweenness scores (financial companies)

Id	**Degree**	**Betw.**	**Clos.**	**Financial**	**Relative betw.**
AEM SPA	0	0.000	1444	0	0
ALITALIA LINEE AEREE ITALIANE SPA	0	0.000	1444	0	0
ALLEANZA ASSICURAZIONI SPA	5	0.064	408	1	0.012
ARNOLDO MONDADORI EDITORE SPA	4	0.049	418	0	0.009
ASSICURAZIONI GENERALI SPA	5	0.018	408	1	0.004
ATLANTIA SPA	9	0.464	395	0	0.089
AUTOGRILL SPA	8	0.617	397	0	0.119
BANCA INTESA SPA	4	0.000	409	1	0
BANCA MONTE DEI PASCHI DI SIENA SPA	1	0.000	1407	1	0
BANCA POPOLARE DI MILANO SCRL	0	0.000	1444	1	0
BANCA POPOLARE ITALIANA - BANCA POPOLARE DI LODI SOCIETA' COOPERATIVA	1	0.000	416	1	0
BANCO POPOLARE DI VERONA E NOVARA SCRL	1	0.000	435	1	0
BULGARI SPA	0	0.000	1444	0	0
CAPITALIA SPA	6	0.135	406	1	0.026

ENEL SPA	1	0.000	435	0	0
ENI SPA	4	0.073	412	0	0.014
FASTWEB SPA	1	0.000	448	0	0
FIAT SPA	5	0.277	413	0	0.053
FINMECCANICA SPA	3	0.024	416	0	0.005
FONDIARIA - SAI SPA	4	0.246	407	1	0.047
GRUPPO EDITORIALE L'ESPRESSO SPA	3	0.239	420	0	0.046
ITALCEMENTI SPA FABBRICHE RIUNITE CEMENTO	6	0.110	401	0	0.021
LOTTOMATICA SPA	1	0.000	425	0	0
LUXOTTICA GROUP SPA	5	0.135	409	0	0.026
MEDIASET SPA	6	0.155	403	0	0.03
MEDIOBANCA SPA	13	0.764	389	1	0.147
MEDIOLANUM SPA	4	0.028	409	1	0.005
PARMALAT SPA	5	0.196	401	0	0.038
PIRELLI & C. SPA	14	1.000	388	0	0.192
SAIPEM SPA	1	0.000	1407	0	0
SANPAOLO IMI SPA	1	0.000	441	1	0
SEAT PAGINE GIALLE SPA	3	0.014	417	0	0.003
SNAM RETE GAS SPA	1	0.000	1407	0	0
TELECOM ITALIA SPA	8	0.289	396	0	0.056
TERNA - RETE ELETTRICA NAZIONALE SPA	2	0.000	412	0	0
UNICREDITO ITALIANO SPA	5	0.246	407	1	0.047
UNIONE DI BANCHE ITALIANE SCPA	5	0.055	407	1	0.011
UNIPOL SPA - COMPAGNIA ASSICURATRICE UNIPOL	1	0.000	1407	1	0

Table 11: Centrality measures and relative betweenness (global network)

6. Discussion

Even though the methodology we are proposing can be applied in different fields, the analysis of organizations is an ideal topic. Organizations are usually structured by departments and they, or the different member of a department, may be identified as the attributes. We can analyse the various centrality measures associated to their members as regards one (or more) specific attributes. Indeed, in order to become more efficient, an organizational structure should clearly take into account of the differences in betweenness scores paying attention to the various organizational schemes of a company.

In general, in organizational studies it is useful to design and analyse network flow in order to assess, and possibly improve, the effectiveness and the efficiency of the organization. In these cases, our method is particularly significant, as it makes it possible to grasp the various structures that can be detected by considering the different attributes. To this end, we have contributed to the structural analysis of a network by identifying relevant subgroups of nodes in the network and analysing their structure. On the one hand, the analysis of the global network structure is useful for understanding the relevant nodes in the structures. On the other hand, analysing the subgroups of the network makes it possible to identify the most relevant various nodes by considering some important attributes (for instance, departments).

We are able with our methodology to complete the structural analysis of the entire network with the structural analysis of the different subgroups identified by attributes. Indeed, we are able to consider the different nodes showing high values of betweenness centrality for both the whole network as well as the various identified subgroups. We can then compare the results in order to analyse the structure of the organizations at an advanced level.

Finally, an interesting result occurs when the different attributes identify the most central nodes of the network; what emerges in this case is that the most central nodes in the subnetworks are also central in the network. This might prove to be essential information, because it states the presence of a significant attribute. In organizational terms, a department or a characteristic of a group of nodes can also be characterized by the attribute.

Conclusions

In this work, a study was done on the problem of the node position in connected subgraphs characterized by a relevant attribute. The methodology we proposed consists of computing a relative centrality value, in order to identify the most central nodes in the subnetwork under examination. Hence, we are able to detect the different nodes that are the most central in the network and/or in the subgraphs. The attribute-based subnetwork has been considered in this work as a network separated by the whole structure and the relative betweenness is computed in the algorithm accordingly.

Applications are possible in various fields, but it can be very useful in organization studies, when assess the efficiency of an organization is required or detecting the information flowing within an organizational structure. The advantage of using this methodology also consists in its immediacy when we need to analyse network partitions, which can be potentially relevant in the network structure.

In this case, it might be important to understand whether these subnetworks correspond to the “core” of the network or they can be identified with the periphery of the network.

Further researches can be developed. This approach can be investigated by computing the relative centrality when the subnetwork is embedded in the network as a whole, as is the case of geographical networks, extending the analysis of this paper. Moreover, other network characteristics (such as communities) can be related to the network structure and our methodology can also be applied to community detection.

Acknoledgement: We want to thank Anna Torriero and Adyemi Sonubi for their careful reading and comments. All error are our own.

Appendix: Algorithm

```
g<-erdos.renyi.game(30, 1/5)
# define the network
p<-rep(1:3,times=10)
# define for each node the attributes
g <- set.vertex.attribute(g, "color", value=p)
# assign to each node their attribute
n<-1:length(p)
V(g)$id<-n
V(g)$value<-p
plot(g)
x11()
# visualize the network (the color of the node is relate their attribute)
be<-betweenness(g)
# compute the betweenness for each node
ris<-data.frame(n,p,be)
# visualize a data-frame with each node, their attribute and their betweenness
######################################################################
s1 <-induced_subgraph(g,V(g)$value=="1")
labels1<-dat$nom[dat$p==1]
plot(s1,vertex.label=labels1)
x11()
# extract and visualize the subnetwork related to attribute “1”
s2 <-induced_subgraph(g,V(g)$value=="2")
labels2<-dat$nom[dat$p==2]
plot(s2,vertex.label=labels2)
```

```
x11()
# extract and visualize the subnetwork related to attribute “2”
s3 <-induced_subgraph(g,V(g)$value=="3")
labels3<-dat$nom[dat$p==3]
plot(s3,vertex.label=labels3)
# extract the subnetwork related to attribute “3”
##########################################################################
# Global Betweenness Tables
ris<-data.frame(p,be)
ris1<-data.frame(as.numeric(V(s1)),betweenness(s1))
ris2<-data.frame(as.numeric(V(s2)),betweenness(s2))
ris3<-data.frame(as.numeric(V(s3)),betweenness(s3))
write.table(ris,'clipboard')
###########################################################################
# Relative Betweenness Tables
riss<-data.frame(p,be,be/sum(be))
riss1<-(betweenness(s1)/sum(betweenness(s1)))
# relative betweenness subnetwork 1
riss2<-(betweenness(s2)/sum(betweenness(s2)))
# relative betweenness subnetwork 2
riss3<-(betweenness(s3)/sum(betweenness(s3)))
# relative betweenness subnetwork 3
tab1<-data.frame(n=dat$nom[dat$p==1],rbe=riss1)
tab2<-data.frame(n=dat$nom[dat$p==2],rbe=riss2)
tab3<-data.frame(n=dat$nom[dat$p==3],rbe=riss3)
tab1
tab2
tab3
```

References

Albert R., Barabasi A. (2002). Statistical mechanics of complex networks, *Reviews of Modern Physics* 74, 47-97

Agryzkov, T., Oliver, J. L., Tortosa, L., Vicent, J. (2014). A new betweenness centrality measure based on an algorithm for ranking the nodes of a network. *Applied Mathematics and Computation*, 244, 467-478.

Barabasi, A. L., Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286 (5439), 509-512.

Bell J. R. (2014). Subgroup centrality measures. *Network Science* 2, 277-297.

Bonacich, P. (1972). Technique for analyzing overlapping membership. In H. Costner, editor, Sociological Methodology, pages 176–185. Jossey-Bass, San Francisco.

Borgatti, S.P., Everett, M.G., Freeman, L.C. (2002). Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.

Brass, D. J. (1984). Being in the right place: a structural analysis of individual influence in an organization. *Administrative Science Quarterly,* 518-539.

Csardi G, Nepusz T (2006): The igraph software package for complex network research, InterJournal, Complex Systems 1695. http://igraph.org

Erdős, P., Rényi, A. (1959). On random graphs. *Publicationes Mathematicae* Debrecen, 6, 290-297.

Dooley, P. C. (1969). The interlocking directorate. *The American Economic Review*, 59(3), 314-323.

Estrada E., Rodríguez-Velázquez J.A., (2005) Subgraph centrality in complex networks, Physical Review E 71, 056103 s2005d

Everett M.G., Borgatti, S. (1999). The centrality of groups and classes. *The Journal of Mathematical Sociology*, 23 (3), 181-201

Everett M. G., Borgatti S., (2005). Extending centrality. *Models and methods in social network analysis* 35.1, 57-76.

Ford L. R., Fulkerson, D. R. (1956). Maximal flow through a network. *Canadian Journal of Mathematics* 8, 399-404.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports* 486(3), 75-174.

Freeman, L. C. (1977). A Set of Measures of Centrality based on betweenness. *Sociometry* 40(1), 35-41.

Freeman, L. C., Borgatti S., White D. R. (1991). Centrality in valued graphs: a measure of betweenness based on network flow. *Social Networks* 13, 141-154.

Girvan M., Newman M.E.J. (2002). Community structure in social and biological networks, *Proc. Natl. Acad. Sci*. USA 99. 7821–7826.

Grassi R., Scapellato R., Stefani S., Torriero A. (2008). Betweenness centrality: extremal values and structural properties. In *Naimzada AK, Stefani S, Torriero A Eds. Networks, Topology and Dynamics, Theory and Applications to Economics and Social Systems. Lecture Notes in Economics and Mathematical Systems,* 613. Springer Heidelberg 161-176.

Gulati, R. (1998). Alliances and networks. *Strategic management journal*, 293-317.

Harary F.(1969). Graph theory, Addison-Wesley, Reading, MA.

Koenig, T., Gogel, R., & Sonquist, J. (1979). Models of the significance of interlocking corporate directorates. *American Journal of Economics and Sociology*, 38(2), 173-186.

Latora, V., Marchiori, M. (2007). A measure of centrality based on network efficiency. *New Journal of Physics* 9 (6), 188.

Leskovec, J., Kleinberg, J., Faloutsos, C. (2007). Graph evolution: densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data* (TKDD), 1(1), 2.

Lux, T. Marchesi, M. (1999) Scaling and criticality in a stochastic multi-agent model of a financial market, *Nature* 397, 498.

Mariolis, P. (1975). Interlocking directorates and control of corporations: The theory of bank control. *Social Science Quarterly*, 425-439.

Mehra, A., Dixon, A. L., Brass, D. J., Robertson, B. (2006). The social network ties of group leaders: Implications for group performance and leader reputation. *Organization Science*, 17(1), 64-79.

Mizruchi, M. S. (1996). What do interlocks do? An analysis, critique, and assessment of research on interlocking directorates. *Annual review of sociology*, 271-298.

Newman M.E.J. (2002). Assortative mixing in networks, *Phys. Rev. Lett.* 89.208701.

Newman, M E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks* 27, 39–54.

Nooteboom, B. (1999). Voice-and exit-based forms of corporate control: Anglo-American, European, and Japanese. *Journal of Economic Issues*, 33(4), 845-860.

Palmer, D. (1983). Broken ties: Interlocking directorates and intercorporate coordination. *Administrative Science Quarterly*, 40-55.

Piepenbrink, A., Gaur, A. S. (2013). Methodological advances in the analysis of bipartite networks: an illustration using board interlocks in Indian firms. *Organizational Research Methods* 16(3), 474-496.

Robins, G., Alexander, M. (2004). Small worlds among interlocking directors: Network structure and distance in bipartite graphs. Computational & Mathematical Organization Theory, 10(1), 69-94.

Smyth P., White S. (2003). Algorithms for estimating relative importance in networks. In: *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '03*, ACM, Washington, D.C., 266–275.

Santella, P., Drago, C., & Polo, A. (2009). The Italian chamber of lords sits on listed company boards: an empirical analysis of Italian listed company boards from 1998 to 2006. Available at SSRN

Schoorman, F. D., Bazerman, M. H., & Atkin, R. S. (1981). Interlocking directorates: A strategy for reducing environmental uncertainty*. Academy of Management Review*, 6(2), 243-251.

Soramaki K., Bech M. L., Arnold J., Glass R.J., Beyeler W. E. (2007). The topology of interbank payment flows. *Physica A: Statistical Mechanics and its Applications*, 39(1), 317-333.

Sparrowe, R. T., Liden, R. C., Wayne, S. J., Kraimer, M. L. (2001). Social networks and the performance of individuals and groups. *Academy of Management Journal*, 316-325.

Kolaczyk E.D., Chua D.B., Barthélemy M. (2009) Group betweenness and co-betweenness: inter-related notions of coalition centrality.

Wang, H., Chang, C. K., Yang, H. I., Chen, Y. (2013). Estimating the relative importance of nodes in social networks. *Journal of information processing*, 21(3), 414-422.

Wasserman, S., Faust, K. (1994). Social Network Analysis: Methods and Applications, Cambridge: Cambridge University Press.

Watts, D.J. Strogatz, S. H., (1998). Collective dynamics of `small-world' networks. *Nature* 393, 440-442