The Comparison of Graph Partitioning Methods on Interlocking Directorates Network in Two-Tier Corporate Governance System

Ondrej Nowak, Ales Kubicek and Jiri Hnilica

Department of Business Economics, University of Economics, Prague W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic

Abstract: Since board members are indirectly nominated as representatives of shareholders, companies sharing a board member may have something in common. This may include ownership structure, business groups or other forms of business cooperation. This paper introduces prerequisites to the analysis of potential overlap between interlocking directorate networks and empirical business environment. This includes designing the interlocking directorate network and the description of methods for its partitioning into smaller regions. The analysis was based on the data sample which included all active joint stock companies in the Czech Republic having more than 50 employees as of September 1, 2012. These criteria were satisfied by 2 906 companies.

Key words: Interlocking Directorates · Network Theory · Two-tier System · Business Groups · Ownership Structure

INTRODUCTION

Corporate governance may be defined as: “Procedures and processes according to which an organization is directed and controlled.” [1]. The governing body does not deal with daily operations of an organization because it is the management’s task. Its role and responsibility is to set up the organization’s strategic direction and oversee, monitor and control the management in the given direction.

Governance in the organization is executed through two basic models: one-tier and two-tier board system (also called unitary and dual board system). In the unitary board system the governing body is called the board of directors which consists of executive and non-executive members. This system is predominant worldwide (e.g. USA, UK).

On the other hand, in the two-tier system there are two separate boards, a supervisory board for the non-executive members and a managing board for the executive members. The managing board as the executive body governs and represents the company, determines the company’s objectives and policies, appoints the company’s management and reports to the supervisory board and to the shareholders at the shareholders meeting. The supervisory board is appointed by shareholders and oversees how the managing board exercises its range of powers and how the business activity of the company is conducted.

The two-tier system can be found exclusively in continental Europe (e.g. Germany, Austria and the Czech Republic). The board system is usually set by law but there are some countries where shareholders can decide which system to adopt (e.g. France, Japan and Slovenia).

For the purpose of this paper, if it is not necessary to distinguish between the executive and the supervisory board type the general term boards will be used. If the discussion requires this distinction, the board type will be specified.

Board interlocks occur when members sit on the boards of more than one company at the same time. This, in theory, increases the cohesion of the business elite, favors cooperation, facilitates coordination of economic activities and permits mutual control [2].

Since board members are indirectly nominated as representatives of shareholders, companies sharing a board member may have something in common. This may include ownership structure or another form of business cooperation. The reasoning behind this claim is subject to ongoing research by the authors and is out of the scope
of this paper. This paper, however, introduces prerequisites to the analysis of the overlap between interlocking directorate networks and the empirical business environment. Furthermore, the design of the network and methods for its partitioning into smaller areas are introduced in this text.

An Introduction to Network Theory: A network (or graph) is a set of items, termed nodes, with connections between them called edges. At this level of abstraction, graphs can be used to represent a vast range of phenomena. Our attention in this paper is restricted to networks derived from the world of corporate boards. For this purpose we adopt the following conventions: our nodes will be of two types – either companies or people. Edges will represent membership of a person in a board of the company. There are also two types of edges – either membership in the managing or the supervisory board.

Next we define the connected component associated with a node: it is that part of the graph consisting of the node itself and all those other nodes that can be reached by paths running along the edges of the graph. A path is simply a sequence of nodes with the property that each consecutive pair in the sequence is connected by an edge. Sometimes it is also useful to think of the path as containing not just the nodes but also the sequence of edges linking these nodes.

The data about board members, however, present an immediate problem: how should one draw a graph to represent it? The issue is that one could treat the company as the basic unit of analysis and form a graph whose nodes would represent companies and whose edges would represent a shared board membership. But alternatively, one could focus on the people and make a (generally much larger) graph whose nodes would represent people and whose edges would represent their shared board memberships. There is no obvious way to choose between these two representations and so most authors simply analyze both.

The ambiguity about representation, however, arises naturally from the structure of the data: there really are two sorts of social entities here, the people and the boards of companies. The network’s edges then represent membership of the former in the latter. The most natural representation of such a network is a graph with two sorts of vertices – one for companies and one for people. The edges of such a network are connecting people with companies in whose boards they sit. Sometimes, this kind of network is called an affiliation network. Including both types of nodes in the graph is called a two-mode projection of the network.

Affiliation Network and its Projections: The network on Figure 1 is called an affiliation network in the two-mode projection, because it includes both types of nodes and their affiliation. Some analysis approaches may, however, find such a network too complicated and will require its simplification. Methods discussed later in this paper focus on companies and for this reason individuals will be removed from the graph.

A sample connected component of the corporate governance network is shown in Figure 1. In the network, two sorts of nodes can be observed: The larger dark nodes represent companies and the smaller brighter nodes represent individuals. Links between nodes show membership of an individual in either executive (darker lines) or supervisory (lighter lines) board.

The connected component on Figure 1 consists of eight companies (nodes A, B, C, D, E, F, G and H) linked by individuals who share memberships in their boards. Structurally important nodes can be identified intuitively from the graph. Company A is located in the center of the network and bridges two tightly knit groups of companies. It shares three individuals with company B (1, 2 and 3), two individuals with company C (1 and 2) and only a single individual with company G (1).

These connections linking companies together through shared membership are called interlocks. As there are twelve unique paths from company A to other companies, we say that company A has 12 interlocks. It is also clear that two companies can be connected by more than one interlock. For example, company A is connected by 12 interlocks to only 6 other companies.

From another point of view, the most important node in the network is individual 1, who is linked to five companies which means that he or she is a member of five boards (he or she holds one executive board membership and four supervisory board memberships). Hence, we say that individual 1 holds five memberships.
The process of removing one class of nodes from the graph results in a so-called one-mode projection shown in Figure 2. Two nodes in Figure 2 are connected by an edge if the companies they represent share at least one member of their board. For example, companies A and D on Figure 1 share two members of their board (nodes 4 and 5) which results in the edge between nodes A and D in Figure 2 [3].

**Graph Partitioning:** Many methods for breaking networks down into a set of tightly knit regions have been developed. Most of them will prove effective for networks with clear divisions into regions. However, since the results of their application may vary due to the logic of the algorithm, it is important to identify general styles that motivate their design.

One group of methods focuses on identifying and removing the bridging edges between densely interconnected regions. Once these links are removed, the network starts to fall apart into large pieces, for which it is possible to identify further bridging edges and the process continues. These methods are called divisive methods of graph partitioning, because they recursively divide up the network. The alternative group of methods starts from the opposite end, focusing on those most tightly connected parts of the network, where nodes are likely to belong to the same region and merge them together. The resulting network will consist of a large number of merged chunks, each containing the seeds of a densely connected region; the process then looks for chunks that can be further merged together and in this way the regions are assembled in the bottom-up fashion. These methods are called agglomerative graph partitioning methods, as they glue nodes together into larger regions.

The process of graph partitioning can be seen as useful when identifying regions that are naturally nested. The larger regions may contain several smaller ones and these may further contain even more densely connected regions nested within them. This is a familiar picture from everyday life where, for example, a global population separates into national groups and these can be further divided into subpopulations within particular local areas within countries.

After their application, most graph partitioning methods will find nested regions. Divisive methods will generally proceed by subsequently removing the bridging edges. Agglomerative methods will arrive at the same result from the opposite direction, by combining the smallest triangles into blocks and then merging the blocks into larger chunks. This is an important starting point to which, in this text, we can further focus on the divisive methods and ignore the agglomerative.

**Betweenness:** Divisive methods proceed by removing bridging edges that span weakly interacting parts of the network. The problem, however, lies in the fact that this approach is not robust enough - and this is for two reasons. Firstly, there are usually several bridges and we need to decide which of them we shall remove first. Secondly, there may be graphs without visible bridges, because every edge belongs to a triangle and yet dividing the chart into smaller areas is still naturally possible. A simple example of such a graph is displayed in Figure 3 where we can see regions of tightly linked nodes 1-5 and 7-11, even though there is no single bridging edge between them.

However, if we think more generally about what bridges do, we can arrive at a notion that is central to the Girvan-Newman method [4, 5]. Bridges are important as they form the shortest path between pairs of nodes in different parts of the network. After the removal of a bridge, new paths between many pairs of nodes have to be identified. Therefore, we may now define the abstract notion of "traffic" in the network and look for edges that carry the most of it. As in the case of key bridges and highway arteries in transportation, we might expect that these edges separate densely connected regions and therefore should be good candidates for removal in the process of dividing the network.

We define the notion of network traffic as follows: For each pair of nodes A and B in the graph that is connected by a path, we may introduce a single unit of "flow" along the edges from A to B (if A and B belong to different components, there is no flow between them). If there is more than one shortest path between A and B, the flow is evenly distributed between them. So if there are k shortest paths between A and B, then each of them carries 1 / k flow.
Now we may define the notion of betweenness of an edge, as the total flow it carries. It represents the number of all possible pairs of nodes whose shortest paths use this edge. For example, in Figure 2 we can calculate the betweenness of edges in the graph as follows:

- Firstly, consider the 5-7 edge. For each node A in the left half of the graph and each node B in the right half of the graph, their full unit flow passes through the 5-7 edge. On the other hand, no flow between pairs of nodes that both lie in the same half of the graph passes through this edge. As a result, the betweenness of the 5-7 edge is $5 \times 5 = 25$.
- The 4-5 edge carries a full unit flow from node 4 to all nodes on its right. This means that the betweenness of this edge is $1 \times 7 = 7$. The same applies to the 7-8 edge.
- The 2-3 edge carries only flow between its own endpoints and therefore its betweenness is 1. The same goes for the 9-10 edge.

The betweenness method thus selected the 5-7 edge as the one that carries the most traffic in the network.

Sociology sometimes focuses on nodes instead of edges, but the definition and the method of calculation remain the same – the betweenness of a node is the total amount of traffic that this node carries in the network. Both edges and nodes with high betweenness occupy critical roles in the network structure. Carrying a large amount of traffic indicates a location of the node at the interface between groups of tightly connected nodes [6].

The Girvan-Newman Method: Edges with high betweenness carry the most traffic along shortest paths. Based on the assumption that these are systematically the most important edges connecting different areas of the network, it is natural to try to remove them first. This idea is at the heart of the Girvan-Newman method, which can be summarized as follows:

- Find the edge with the highest betweenness (or more edges, if there is a tie) and remove these edges from the graph. This can divide the graph into several components. If so, this is the first level of regions in the graph partitioning.
- Now we have to recalculate all betweenness values and again remove the edge or edges with the highest betweenness. This may again break some of the existing components into smaller components and if so, these are the regions nested into larger regions.
- We proceed in this way as long as there are any edges in the graph. At each step, we recalculate all betweenness values and remove the edge or edges with the highest value.

In this matter, the graph falls apart first into large pieces and then, at the lower levels, the method naturally reveals the structure of nested tightly interconnected regions. Figure 4 shows how the method operates on the graph from Figure 3. It is clearly visible how by removing edges, the graph breaks from larger regions down into smaller ones.

The sequence of steps in Figure 4 reveals a few interesting points about how the method works:

- When we calculate the betweenness in the first step, the 5-7 edge carries all the flow between nodes 1-5 and 7-11 and its betweenness is 25. On the other hand, the 5-6 edge carries only flow from node 6 to each of the nodes 1-5 and its betweenness is 5 (the same applies to the 6-7 edge).
- After removing the 5-7 edge, it is necessary to recalculate all betweenness values. All 25 units of flow that were carried by the removed edge have shifted to the path through nodes 5, 6 and 7. The betweenness of edges 5-6 and 6-7 therefore increased to $5 + 25 = 30$, which is the reason for removing the two edges in the next step.
In the first presentation of their method, Girvan and Newman demonstrated its effectiveness by breaking down several real networks into intuitively reasonable sets of regions [4, 5].

**The Empirical Network of Interlocking Directorates:**
The data sample included all active joint stock companies in the Czech Republic having more than 50 employees as of September 1, 2012. These criteria were satisfied by 2,906 companies. For these companies all the active members of both their executive and supervisory board were identified and exported into a spreadsheet. More than 20,000 board memberships were identified with 17,699 individuals who were assigned a unique randomly generated ID to be used for further anonymous analysis.

It is a common practice that data sample for an empirical investigation of corporate governance consists of listed companies. There are, however, only 28 listed companies on the Prague Stock Exchange which would not be considered a representative data sample. Nevertheless, there are no obligations related to corporate governance requested by the Prague Stock Exchange. Hence, it is fully in the competence of companies as to whether they comply with good governance practices and it is at their own discretion to follow the Czech Code of Corporate Governance based on the OECD Principles.

For investigation purposes, publicly disclosed interlocking directorates network information from two local information sources in the Czech Republic was used: Official government website www.justice.cz and database Magnus maintained by CEKIA.

**An Empirical Network:** Based on the logic described in the introduction of this paper, the data set has been converted into an interlocking directorate network, which will be analyzed in the one-mode projection (i.e. all nodes representing individuals will be removed from the graph so that it contains only companies). The edge between two nodes in a one-mode projection represents shared membership of an individual in boards of the two companies.

The resulting network consists of many components. In order to maintain the simplicity of the discussion, only the largest (main) component of the network will further be used. The methods introduced in the following text are naturally applicable to any network component consisting of more than one node.

The largest component of the interlocking directorates network in the Czech Republic is shown in Figure 5. It consists of 192 companies. The saturation of the node color reflects the betweenness value. Darker nodes exhibit relatively higher betweenness. These nodes are expected to be located at the ends of edges that span areas of separate tightly connected regions.

Tightly knit regions of nodes connected together by bridges are clearly visible in Figure 5. As discussed in the introduction, these regions might reflect real business groups. The analysis of the potential overlap between the interlocking directorates network and empirical business structures requires the isolation of these regions. In the following text, two approaches to their isolation will be applied to the network.

**Partitioning of the Main Component Using the Girvan-Newman Method:** The Girvan-Newman method was employed first. After three steps, three edges with the greatest value of betweenness were removed. This led to the disintegration of the chart into four regions as shown in Figure 6. It is clearly visible that the three deleted edges were bridges with end nodes with the highest betweenness as highlighted in Figure 5.

Intuitively, it would also be possible to further break down the dominant region a) in Figure 6 in the middle. The region could be split in the area around the black node highlighted by the arrow in the center of Figure 5. But the node is tied to the rest of the network by 15 edges among which the betweenness spreads. For this reason the region holds together.

Removing the three edges therefore represents the optimal number of steps for a sufficient breakdown of the network. Further steps of the Girvan-Newman methods lead to the separation of only small groups of nodes.
Fig. 6: Break down of the main component from Figure 5 after the application of the Girvan-Newman method

Let's take a closer look at the four newly created subcomponents in Figure 6. We may now try to describe them from the real business environment point of view. We start with region a) consisting of 112 nodes, which represented the dominant cluster of the main component in the original network in Figure 5. The apparent dominance here is provided by companies belonging to the business group around the Czech agricultural tycoon Andrej Babis. In fact, the center of the region is occupied by Agrofert Holding, the parent company of the group.

Region b) consists of companies belonging to the J&T group, operating mainly in the energy and utilities sector. Central companies to the group were United Energy and Opatovice power plant.

Region c) consists of 24 companies with the highest representation of the coal industry having Czech Coal in its center. The coal industry in the Czech Republic is dominated by a few business groups including the above mentioned J&T and the national enterprise CEZ.

Analysis of 36 nodes in region d) suggests the superiority of the CEZ Group, with the parent company CEZ in the center. The region also includes Dalkia group, a number of power plants and few other national enterprises such as MERO, CEPRO and Czech Airlines.

Partitioning of the Main Component Using the Membership Type: Since interlocking directorate networks consist of companies linked together by the shared membership of the executive and the supervisory board, we can use this distinction to divide the main component by removing all edges representing the membership in the supervisory board. In this way, the network can be expected to break into clusters of companies with potentially close business ties. The reason for this is that these companies share only members of the executive board occupied by corporate executives closely associated with the company.

The appearance of the main components after the removal of all links representing the supervisory board membership is shown in Figure 7. The component disintegrated into many smaller clusters. Compared to Figure 6, the largest region a) was again identified in the middle, while having Agrofert Holding still in the central position. However, this time the region is much smaller and two other regions e) and f) span off. It is worth mentioning that this time region e) separated as was intuitively expected.

Regions b) and c) are still visible, even though they contain far fewer nodes compared to the layout generated by the Girvan-Newman method. On the other hand, region d), consisting of companies from CEZ Group and other national enterprises, collapsed completely.

CONCLUSION

Two methods for graph partitioning were applied to the interlocking directorate network in the two-tier corporate governance board system in the Czech Republic. The Girvan-Newman method led to larger compact groups of companies, while the membership type approach resulted in much smaller but more compact groups.

It was interesting to distinguish the structural significance of the membership type within interlocking directorate networks. Shared membership in the executive board often occurs between companies with close business ties. Groups of companies within regions generated by the membership type graph partitioning method therefore much closer resemble business groups.

On the other hand, shared membership in supervisory boards seems to have the character of bridges spanning different areas of the business environment. This is probably the major cause of the network disintegration into separate business groups after they are removed.

This conclusion is not surprising, since in a two-tier system the supervisory board should be composed of individuals with no financial or personal interest in the company. Only this way can they help shareholders to control the management of the company.
REFERENCES


