Clubs of Clubs: Fragmentation in the Network of Intergovernmental Organizations

Yonatan Lupu
Department of Political Science, George Washington University
ylupu@gwu.edu

Brian Greenhill
Department of Government, Dartmouth College
brian.d.greenhill@dartmouth.edu

July 26, 2016

Appendix A: Modularity Maximization

In this appendix, we explain the formal definition of modularity and the algorithm we use to maximize it. Assume we are given a graph (or network) $G = (V, E)$ with n nodes (or vertices) V and m links (or edges) E, with weights w on the links. The edges in this network are directed, meaning they are asymmetric (i.e., state A's trade dependence on state B may be different from state B's trade dependence on A). The $n \times n$ adjacency matrix of the graph G can then be defined as $A_{ij} = w_{ij}$ whenever there is a directed link (ij), and $A_{ij} = 0$ otherwise. The incoming degree of a node i (i.e., the number of nodes that connect to node i) is denoted by $k_i^{in} = \sum_j A_{ji}$ and the outgoing degree (i.e., the number of nodes to which node i connects) by $k_i^{out} = \sum_j A_{ij}$.

Assume each node i is assigned to a cluster σ_i. The modularity of such a partition can

be defined as:

$$Q(\{\sigma\}) = \sum_{ij}(A_{ij} - p_{ij})\delta(\sigma_i, \sigma_j), \tag{1}$$

where $\delta(\sigma_i, \sigma_j) = 1$ if and only if $\sigma_i = \sigma_j$, i.e., where i and j are in the same cluster (Newman and Girvan, 2004; Reichardt and Bornholdt, 2006). Let p_{ij} be an expectation value of an ij link that is taken to be:

$$p_{ij} = \frac{k_i^{out} k_j^{in}}{m}. \tag{2}$$

for a directed network (Leicht and Newman, 2008). This p_{ij} is used to compare a random null model to the empirical trade network, in this case with similar degrees.

Consider the graph $G = (V, E_1, E_2, \ldots, E_T)$ where $E_1, \ldots, E_T$ the represents the links at time $1, \ldots, T$ between the nodes V. We denote by $A_{t,ij}$ the associated adjacency matrix for each ij link at time t. The modularity formula as given in Equation 1 can then be extended slightly to yield:

$$Q(\{\sigma\}) = \sum_{t}\sum_{ij}(A_{t,ij} - p_{t,ij})\delta(\sigma_{t,i}, \sigma_{t,j}), \tag{3}$$

where $\sigma_{t,i}$ now represents the cluster of node i at time t.

Over the past few years, many algorithms for optimizing modularity have been suggested (for an overview see Fortunato, 2010; Porter, Onnela and Mucha, 2009). Because the problem is NP-hard (Brandes et al., 2006) it is unlikely that there is an efficient algorithm to solve the optimization problem perfectly. There are, however, some algorithms that are both efficient (i.e., they run in almost linear time) and effective (i.e., they can correctly identify clusters in test settings) (Lancichinetti, Fortunato and Radicchi, 2008; Lancichinetti and Fortunato, 2009). The so-called Louvain method developed by Blondel et al. (2008) is especially suitable for optimizing modularity. In brief, the Louvain method works as follows. We start out by assigning each node to its own cluster, such that at the start there are as many clusters as there are nodes. We loop (randomly) over all nodes and add them to

a cluster that increases the modularity as much as possible. Then we form a new graph in which each node represents the clusters found at the previous level, with links between these new nodes representing the weights between each cluster in the old graph. In this way, smaller and smaller graphs are obtained, with nodes representing clusters (and possibly sub-clusters). The algorithm ends when modularity can no longer be increased.

Formally, the algorithm works by first removing node i from its cluster, and then calculating the effect on the modularity measure of adding node i to a cluster–possibly the same one. The effect on modularity of putting node i in cluster r in time t can be written as:

$$\Delta Q(\sigma_{t,i} \to r) = \sum_t (m_{t,ir} + m_{t,ri}) - ([m_{t,ir}] + [m_{t,ri}]), \tag{4}$$

where $m_{t,ir} = \sum_j A_{t,ij}\delta(\sigma_{t,j}, r)$ denotes the total weight from node i to cluster r, with $m_{t,ri}$ defined similarly, and $[m_{t,ir}] = \sum_j p_{t,ij}\delta(\sigma_{t,j}, r)$ the expected weight from i to r with again $[m_{t,ri}]$ defined similarly. Each node is then added to the cluster for which this effect on modularity is maximal.

After we have completed the first level, we aggregate the clusters into nodes for a new graph, and define the weight of the links between these new nodes dependent on the clusters. Considering clusters r and s, the total weight from cluster r to s can be written as $m_{t,rs} = \sum_{ij} A_{t,ij}\delta(\sigma_{t,i}, r)\delta(\sigma_{t,j}, s)$. Using this as the weight of the link between node r and s in the new network, the expected value of this link can then be written as:

$$p_{t,rs} = \frac{k_{t,r}^{out} k_{t,s}^{in}}{m_t} \tag{5}$$

$$= \sum_{ij} A_{t,ij}\delta(\sigma_{t,i}, r) \sum_{ij} A_{t,ij}\delta(\sigma_{t,j}, s)\frac{1}{m_t}, \tag{6}$$

which is exactly the expected value of the links between clusters r and s in the old network. Hence, joining nodes r and s in the new network corresponds to joining clusters r and s in

the old network. Doing this for all types of links then gives us a correct new network, upon which we can iteratively apply the method described above. We stop the procedure if we can no longer increase modularity.

As noted above, multiple algorithms have been proposed for maximizing modularity. Some algorithms may be more suitable for a given network than others, i.e., some may perform better at maximizing modularity. Because the purpose of these algorithms is to identify the network partition that maximizes modularity, if different algorithms identify different such partitions, one should choose the result with the largest modularity. We have used the Louvain algorithm because, for our data, it detected partitions with larger levels of modularity than other algorithms. For each year, Figure 1 provides the maximum modularity value generated by the Louvain algorithm and compares this value with those produced by the leading eigenvector (Newman, 2006), fast and greedy (Clauset, Newman and Moore, 2004), walktrap (Pons and Latapy, 2005), spinglass (Reichardt and Bornholdt, 2006), and edge betweenness (Newman and Girvan, 2004) algorithms. For each year, the Louvain method performs better than all the other algorithms at maximizing modularity.

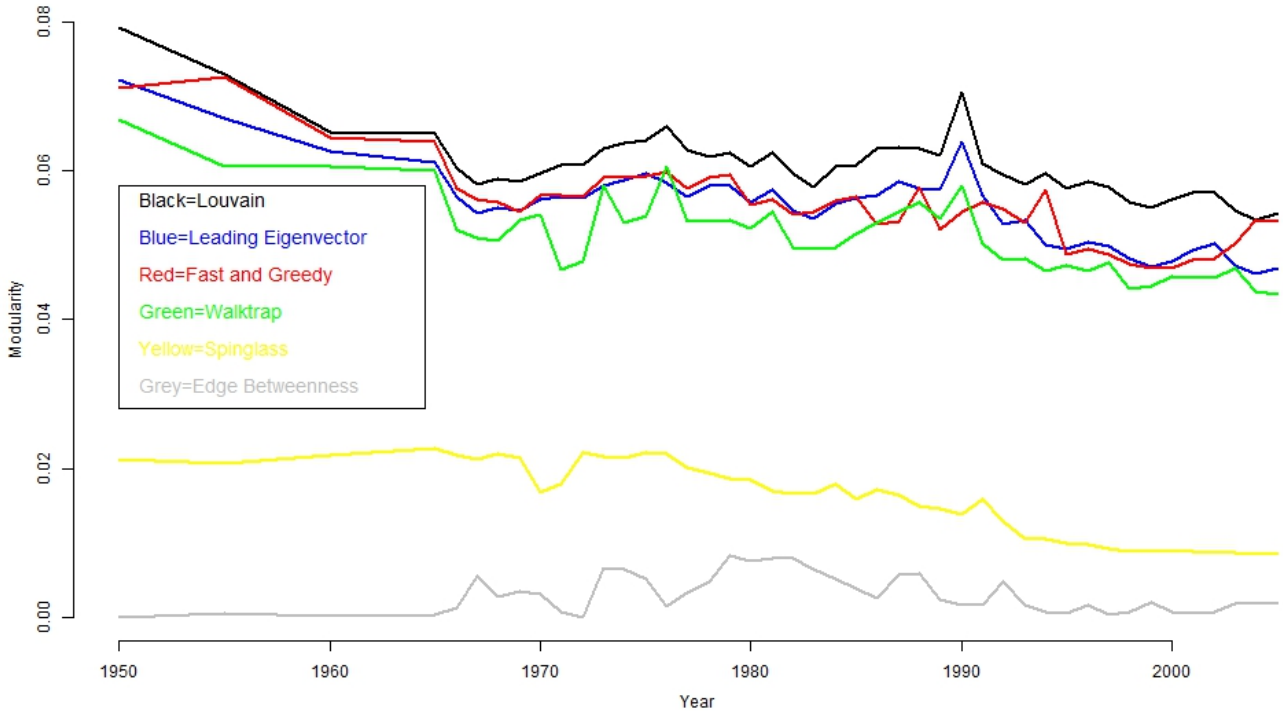

Figure 1: Comparison of Modularity Maximization by Competing Algorithms

Appendix B: Bipartite Network Model

There are at least two ways of modeling the IGO membership data from a network perspective. The approach used in our paper and in the bulk of the existing literature (e.g., Hafner-Burton and Montgomery, 2006; Dorussen and Ward, 2008; Cao, 2009, 2010, 2012; Kinne, 2013) uses a one-mode approach in which states are the nodes in the network. The edge (or tie) between a pair of nodes is typically given a weight based on the number of IGOs to which both states belong.

An alternative approach would be a two-mode approach in which both states and IGOs are modeled as nodes. In such an approach, an edge exists if a particular state belongs to a particular IGO. Edges between states are not possible, and edges between IGOs are not possible. Edges in this approach either exist or do not exist, so they are not weighted. This is the approach used by Beckfield (2010).

While we adopt the former approach in order to be consistent with the bulk of the literature on the IGO network, in this Appendix we test the robustness of our key result to the latter approach. That is, we test whether modularity has decreased over time when we model the network as a two-mode (or bipartite) network. We do so using the algorithm provided by Dormann and Strauss (2014). Figure 2 shows the levels of modularity over time detected using this alternative network model. As the figure shows, our key result – the decrease in the level of modularity over time – is robust to this alternative network model.

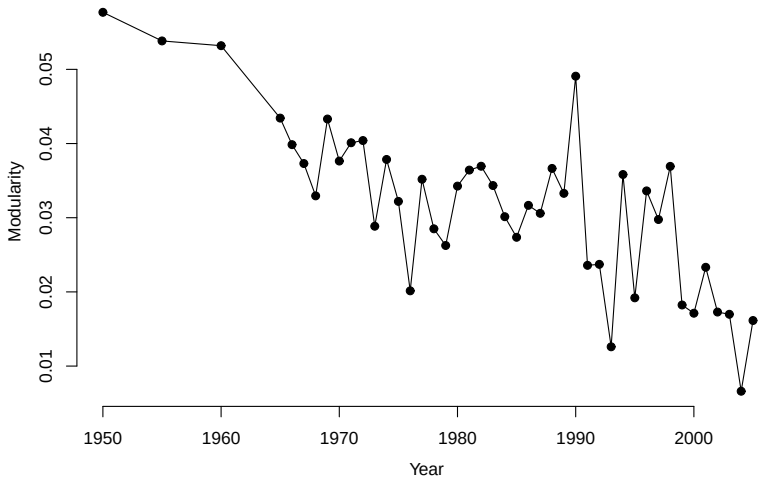

Figure 2: Bipartite Network: Trends in Modularity Scores , 1950-2005.

Appendix C: Economics IGOs Removed

Many IGOs address specifically economic affairs, so it may be the case that joining such IGOs (and, in turn, membership in IGO clusters) is driven in part by states' levels of economic development. To test whether economics IGOs drive our key result, we analyzed models of the IGO network that exclude IGOs that primarily serve economic functions. Specifically, we excluded IGOs that primarily serve either standardization and harmonization, economic cooperation and development, or industry-specific functions (for a total of 105 IGOs excluded).

Figure 2 shows the levels of modularity over time detected using the network models that exclude these IGOs. As the figure shows, the decrease in the level of modularity over time is robust to this alternative specification.

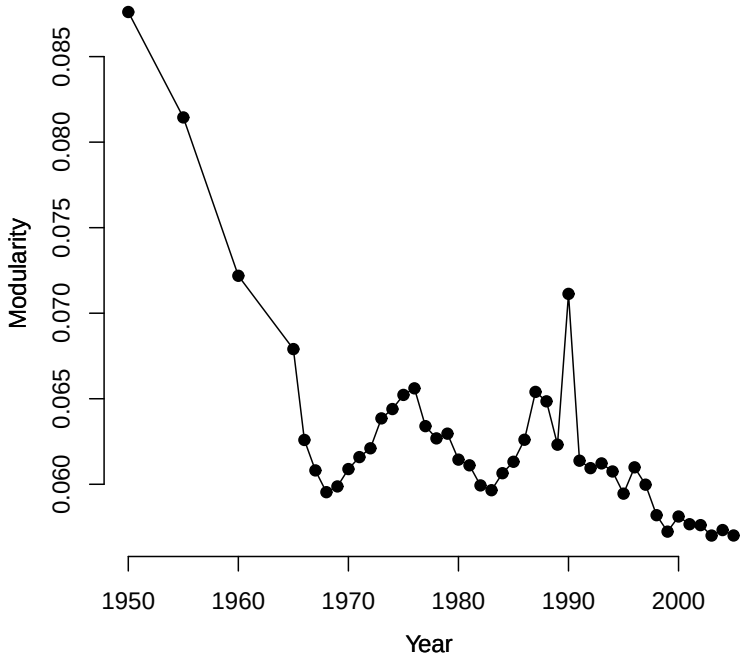

Figure 3: Economic IGOs Excluded: Trends in Modularity Scores, 1950-2005.

References

Beckfield, Jason. 2010. "The Social Structure of the World Polity." *American Journal of Sociology* 115:1018–68.

Blondel, Vincent D., Jean-Loup Guillaume, Renaud Lambiotte and Etienne Lefebvre. 2008. "Fast Unfolding of Communities in Large Networks." *Journal of Statistical Mechanics: Theory and Experiment* 2008(10):P10008.

Brandes, Ulrik, D. Delling, M. Gaertler, Robert Görke and M. Hoefer. 2006. "Maximizing Modularity is Hard." *arXiv* .

Cao, Xun. 2009. "Networks of Intergovernmental Organizations and Convergence in Domestic Economic Policies." *International Studies Quarterly* 53(4):1095–1130.

Cao, Xun. 2010. "Networks as Channels of Policy Diffusion: Explaining Worldwide Changes in Capital Taxation, 1998–2006." *International Studies Quarterly* 54(3):823–854.

Cao, Xun. 2012. "Global Networks and Domestic Policy Convergence: A Network Explanation of Policy Changes." *World Politics* 64(3):375–425.

Clauset, Aaron, Mark E.J. Newman and Cristopher Moore. 2004. "Finding community structure in very large networks." *Physical review E* 70(6):066111.

Dormann, Carsten F. and Rouven Strauss. 2014. "A method for detecting modules in quantitative bipartite networks." *Methods in Ecology and Evolution* 5(1):90–98.

Dorussen, Han and Hugh Ward. 2008. "Intergovernmental Organizations and the Kantian Peace: A Network Perspective." *Journal of Conflict Resolution* 52:189–212.

Fortunato, Santo. 2010. "Community Detection in Graphs." *Physics Reports* .

Hafner-Burton, Emilie M. and Alexander H. Montgomery. 2006. "Power Positions : International Organizations, Social Networks, and Conflict." *Journal of Conflict Resolution* 50:3–27.

Kinne, Brandon J. 2013. "IGO membership, network convergence, and credible signaling in militarized disputes." *Journal of Peace Research* p. 0022343313498615.

Lancichinetti, A. and S. Fortunato. 2009. "Community Detection Algorithms: A Comparative Analysis." *Physical Review E* .

Lancichinetti, A, Santo Fortunato and Filippo Radicchi. 2008. "New Benchmark in Community Detection." *Physical Review E* 78:046110.

Leicht, E A and M E J Newman. 2008. "Community Structure in Directed Networks." *Physical Review Letters* 100(11):118703+.

Newman, Mark E.J. 2006. "Finding community structure in networks using the eigenvectors of matrices." *Physical review E* 74(3):036104.

Newman, Mark E.J. and Michelle Girvan. 2004. "Finding and evaluating community structure in networks." *Physical review E* 69(2):026113.

Pons, Pascal and Matthieu Latapy. 2005. Computing communities in large networks using random walks. In *Computer and Information Sciences-ISCIS 2005*. Springer pp. 284–293.

Porter, Mason A., Jukka-Pekka Onnela and Peter J. Mucha. 2009. "Communities in Networks." *Notices of the American Mathematical Society* 56(9):1082–1097.

Reichardt, Jorg and Stefan Bornholdt. 2006. "Statistical mechanics of community detection." *Physical Review E* 74(1):016110.