

Comparative Analysis of Legislative Complex Networks

Benjamin Lai, Cate Jiang, and Jonathan Chen

<https://github.com/jowch/CSE-416-Final-Project>

December 12, 2018

1 Introduction

In the United States, politics has increasingly become a point of conversation due to change in presidency and new, controversial bills being signed into law. Cosponsorship provides a rich network between legislators that can be used to explain the political effects and efficiency of legislative processes. We propose to analyze the networks of different legislative bodies in select North American and European countries to determine how legislators interact amongst each other to cosponsor bills. We aim to answer questions such as which countries' legislators are most connected, how legislative parties and other attributes play into cosponsorship data, and what that means in terms of government performance and productivity. To perform our research, we will use cosponsorship data from different countries followed by analysis and comparison of statistics and centrality measures.

According to a study conducted by James Folwer from the Department of Political Science at the University of California, Davis, there are several trends that have been observed among legislators. In the United States, biennial elections cause members of the U.S. House and Senate to change every two years, but the trends remain stable. Firstly, party affiliation and similar values and ideals often lead to increased possibility of cosponsorship. Second, average number of cosponsors per legislator reflects the degree that the individual is integrated into their political network [?]. Because these studies were performed solely within U.S. Congression,

we propose to analyze legislators from different first-world countries to determine if these trends would be observed in other political environments as well.

2 Materials and Methods

2.1 Data Collection and Pre-Processing

We obtained cosponsorship data for the countries: Austria, Denmark, France, Sweden, Switzerland, United Kingdom, and the United States. The European networks were produced by a French researcher's website [?]. The United Kingdom Parliament was scraped from Early Day Motions [?], a web portal listing bills discussed. The United States Congress data was scraped using a public domain tool [?]. The raw data was parsed and saved as `graphml` files.

2.2 Data Analysis

Our analysis was performed using the library `graph`, which provides Python bindings to powerful abstractions and parallel algorithms for processing graphs. In addition, we used `numpy`, `scipy`, `pandas`, `seaborn`, and `matplotlib` to perform the comparative analysis.

We wrote a summary function that returns a summary of graph statistics with a given graph. These statistics include: number of edges, number of vertices, average clustering coefficient, average degree, average excess degree, average path length, size of largest connected component, and number of connected components. We then wrote a script that produces a summary output in a csv file for every

country for every year.

An important attribute of our generated networks that we needed to measure was assortativity. Each graph had various vertex properties to perform assortativity measures on, but we decided to use ‘party’ to determine whether governments cooperate on a partison or bipartisan basis. We mapped each ‘party’ to a unique ID, and joined this party enumeration property map with the original vertex properties. We then produced graphs for every country across each year, with nodes colored based off their respective enumerated ‘party’ property map.

Betweenness, closeness, and eigenvector centrality measures were also performed on our data to identify influential legislators. The graphed vertices and edges were scaled and colored with respect to its corresponding centrality measure. Vertices and edges that are yellow and larger in size reflect greater importance, whilst those that are blue and smaller in size reflect lesser importance. For the generated betweenness graphs, the top three percent of legislators with the highest betweenness centralities were colored yellow and enlarged.

3 Results and Discussion

3.1 Country-level Results

We will now go through the summary results for each country observed. Every country we collected data on had temporal data except for Denmark and the UK, which was due to the structure of the data resource. However, we still produced the same summary statistics and produced graphs whenever possible.

For Austria, the number of vertices and edges did not fluctuate between the years, and maintained average values of 63.1 and 63.0 respectively. As expected, average degree also stayed relatively stable. Excess degree decreased dramatically from 1995 to 1999, dropping from 4.5 to 2, which is similar to the average degree. This happens to coincide with Austria joining the European Union in 1995. Although it is unclear whether there is a correlation between these two events, it can be noted that the drop in excess degree close to average degree means that the degree of individuals are roughly the same as their neighbors. This means that there was on

average a sharp decrease in hub behavior.

The French legislative bodies had a steady increase of verticies, starting with less than 300 verticies in 1986 to 500 today. The number of edges peaked during 2002-2007 at 16,000. The average distance has not decreased dramatically, which is surprising given the increase in verticies.

Switzerland’s graph statistics have been stable over the years observed. The country’s clustering coefficient remains around 0.5, its number of verticies around 130, and its number of edges around 1750. According to BBC, Switzerland’s neutrality and political stability has allowed it to become one of the world’s wealthiest countries [?]. This knowledge helps explain why Switzerland’s legislative data has barely changed over the last decade.

There was an interesting change in the number of cosponsorships across Swedish parties between the years 2014-2018. The number of vertices across years consistently hover around 350, whereas the number of edges dropped from 4000 in 1988-1991 to 2000 in 2014-2018. This result happens to coincide with the 2014 Swedish general election, in which the Social Democrats declared that they would not work with the Swedish Democrats. This decrease in edges across parties is evident when comparing graphs from the years 1988-1991 and 2014-2018, seen in Figures 1a and 1b

The United States has an extremely consistent number of edges (around 100,000), vertices (a little over 500), clustering coefficient (almost 1.0), average degree (around 350), and average excess degree (around 400). This shows that despite changing some members every two years, the legislators and their interactions are relatively stable. Most notably, the US legislative bodies have highest completeness (67%; Figure 2a) amongst the examined graphs, where completeness is the ratio between the number of edges in a graph and the number of edges in a complete graph with the same number of vertices.

For the UK data, the average vertices is around 675, which is greater than the average number of edges (663). This means that generally, each node has slightly less than one edge. The average excess degree is greater than 200, which is more than 100 times the average degree. This can also be explained when observing a subset of the betweenness graph, as seen in Figure 2b. The majority of the UK betweenness graph has a single node that is

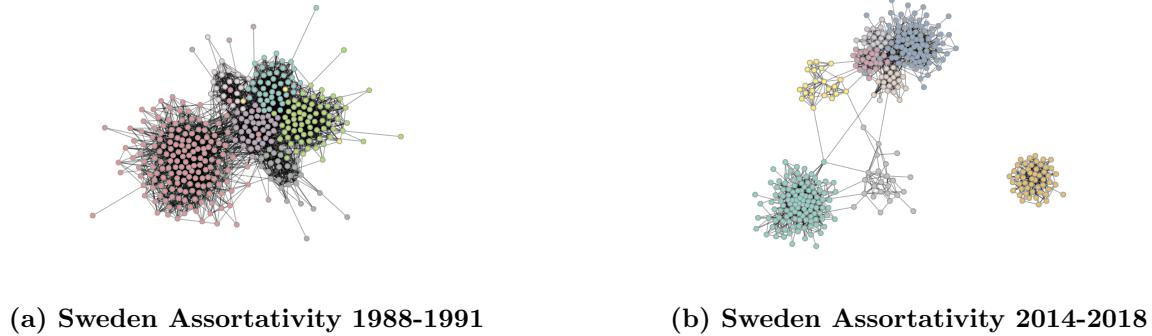


Figure 1.

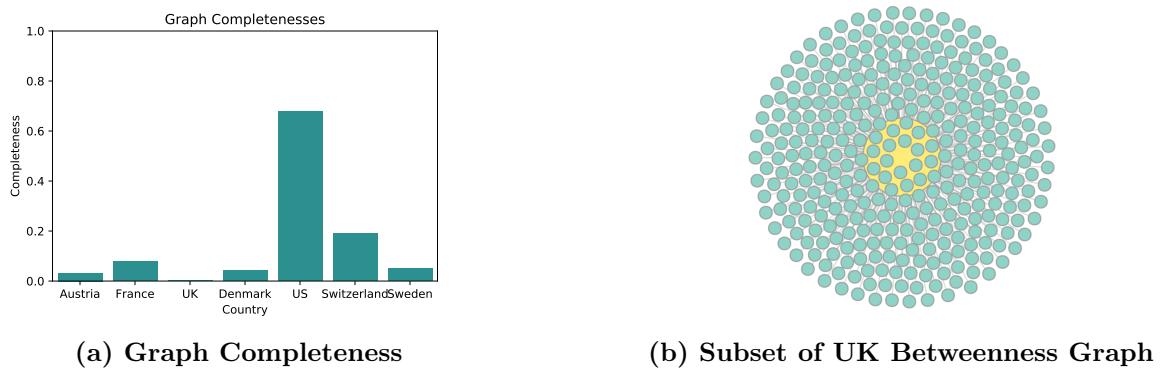


Figure 2.

surrounded by many other nodes.

In Denmark, there are two categories that have the highest number of vertices: Economy and Justice. According to Forbes in December 2017, Denmark is known for its extremely efficient economy and equitable distribution of income. It is also currently undergoing an economic expansion. Denmark is rated #1 for least corruption [?]. Given this information, it is not surprising that Denmark would have the most legislators assigned to Economy and Justice.

3.2 Across-Country Results

The next step to our analysis is to compare results between countries. When comparing the average number of edges and vertices for each country, the UK has the largest number of vertices but the fewest number of edges, signifying that the country has a very low average degree. This property can be observed in Figure 3a.

However, we can also compare the average de-

gree and excess degree across countries. Despite the low average degree, the UK also has the second largest average excess degree. This means that although each vertex is not connected to many other vertices, they are on average connected to relatively important individuals. On the other hand, the US has the highest average degree and the highest average excess degree. This can be observed in the Figure 3b

We notice that the graph structure of the US - much like that of the UK - is separated into exclusive clusters. But unlike the UK, the US depicts an extreme partisanship structure in Figure 4a.

When compared to other countries, the US has a much larger average number of edges and has the second smallest average distance. These results are not surprising. Since legislators in the US are highly connected, the distance between any two random people in the same cluster is much less compared to that of other countries. We also observe that the generated graphs reflect the underlying structure of governance. For example, from the France graph in

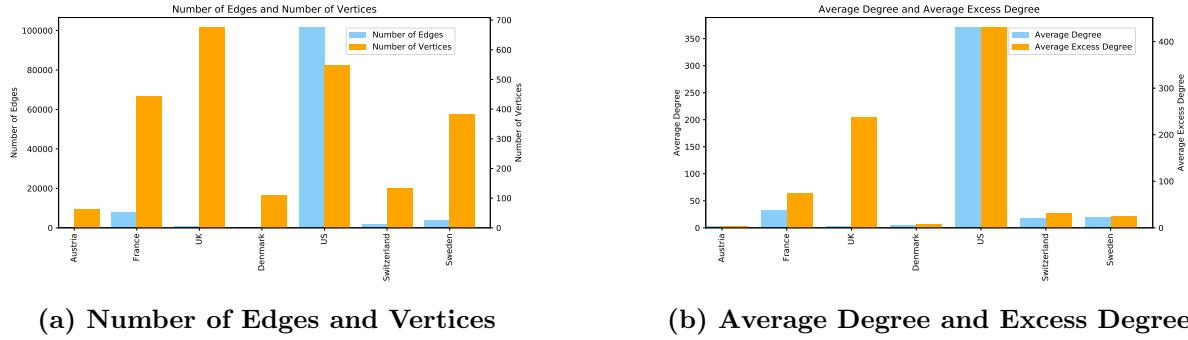


Figure 3.

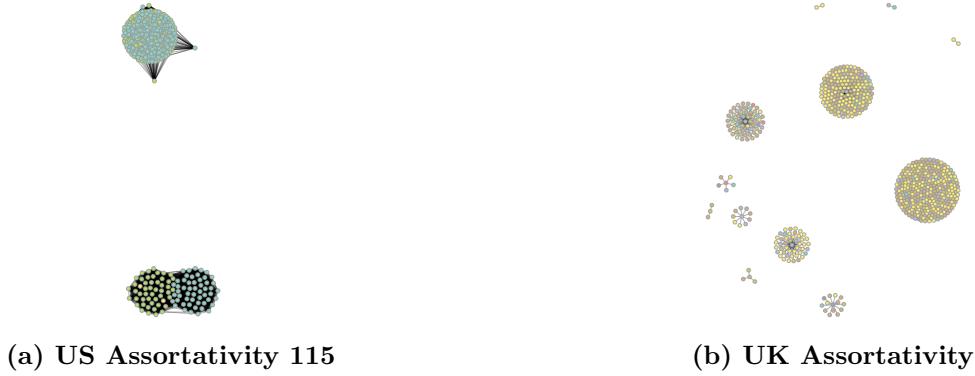


Figure 4.

Figure 5a we notice that the pink component has a single central legislator, and one can thus make an educated guess that the cluster represents the French Socialist Party.

Assortativity measures can be seen in Figure 5b. Assortativity values are most positive for Sweden (around 0.8) and most negative for Austria (around -0.45). Positive assortativity measures means that nodes tend to associate more with nodes with similar attributes, while negative assortativity means that nodes associate with nodes with different attributes. The net US assortativity value is very close to 0, which indicates that there is no overall preference in either direction.

We also expect that each country should exhibit power-law distributions, as there should exist hubs within legislative members with very high degree centralities. This turns out to be the case for smaller less densely connected graphs such as Sweden in Figure 6b. However, this does not hold true for largely connected graphs such as the US, as seen in Figure 6a. The US graph is almost a completely connected graph, where fully connected graphs have

$n(n - 1)/2$ edges. There are roughly 100,000 edges and 552 vertices in our US graph. In this case a complete graph would have 152,076 edges, so the US data holds about 67% of expected edges when compared to a fully connected graph.

4 Discussion

For our project, we use legislative cosponsorship networks across the US and selected European countries to try to infer social relationships in government that may influence legislative behavior. How well officials work together has a direct impact on the efficiency and productivity of government. The dynamic social relationships between legislators - of forging new relationships and maintaining older established relationships - matter. Moreover, changes in network properties can hint at past and ongoing events that impact legislative cosponsorships.

It should be noted that network properties - such as path length and clustering coefficient - of

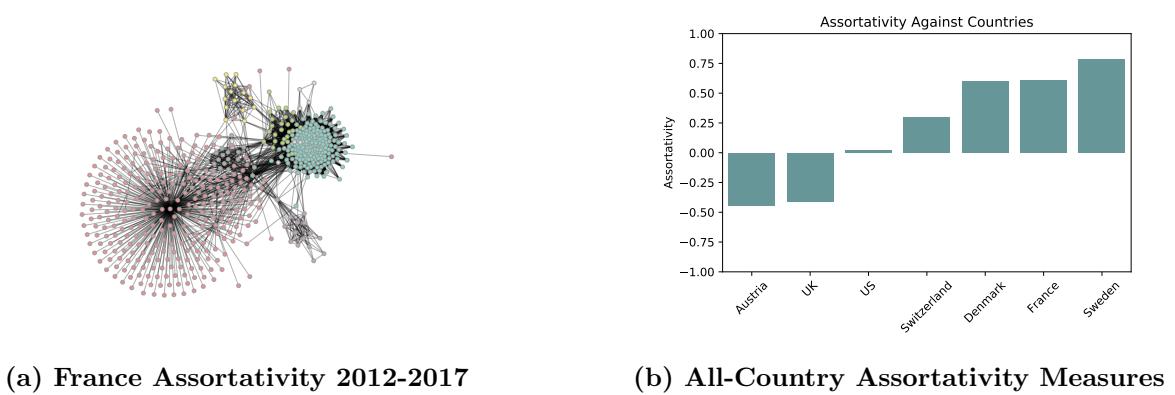


Figure 5.

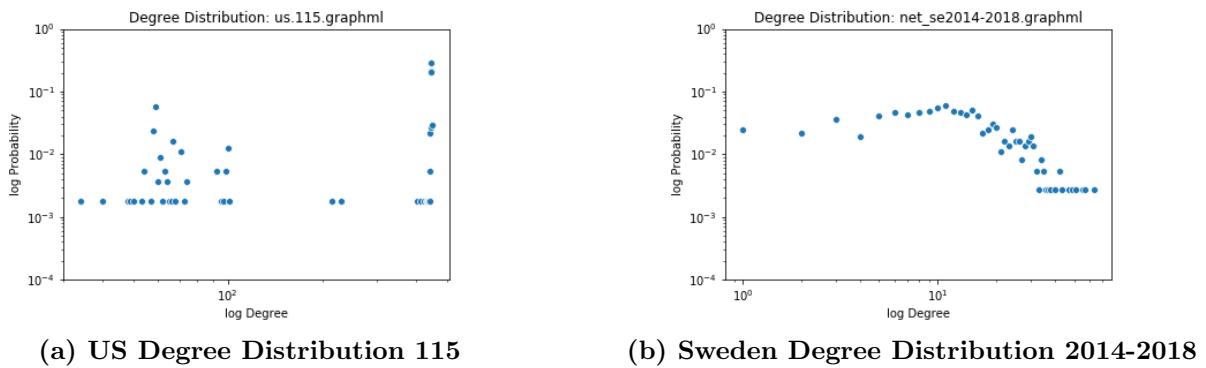


Figure 6.

major developed countries like the US and France tend to show little fluctuation over time. These consistencies can suggest that the changes in governance do not have major influences on the quality of legislative processes, which is a positive sign because these developed countries have proven working systems of governance, democracy, and law.

We have uncovered some intriguing results. First, we notice that the average distance tends to increase as the world gets smaller. One interpretation of this result is that members tend to be more isolated and fewer are willing to cooperate outside of their local network. As the world gets larger, the communities that define and maintain social relationships grow. As communication between legislators improves, links become increasingly interconnected, and the relative distance between any two given legislators decreases. This underlying phenomenon is reflected in the comparison of clustering coefficients across countries. For example, small countries like Austria or Denmark have low clustering coefficients of 0 and 0.2 respectively, whereas a

large country like the US has a clustering coefficient of nearly 1.

We also notice extreme partisanship in the US senate, which is expected given the party division and clashes between the Republicans and Democrats. The two components of the assortativity graph represent the house and the senate. The house and the senate differ from an institutional design standpoint and have contrasting social structures, and so we would expect the behavior in these two components to exhibit unique characteristics as well. Although we did not conduct any direct tests, the house and senate should differ in quantities such as clustering coefficient and average path length.

Lastly, we examined the degree distributions for every graph, and hypothesized that they would follow power-law distributions. As it turns out, this does not hold true for large densely connected graphs and non scale-free networks.