Congressional Collaboration Within and Between Terms

Alex R. Goodman
B.S. Candidate in Symbolic Systems
Stanford University
alexgood@stanford.edu

Emma L. Mclean
B.S. Candidate in Symbolic Systems
Stanford University
emclean@stanford.edu

Austin J. Zambito-Valente
B.A.S. Candidate in Computer Science + Music
Stanford University
azv04@stanford.edu

Abstract

There is a common conception that politics in the United States has become more polarized over the course of the past generation, and especially during the reigns of our past few presidents. We use network science on House of Representatives vote data since the Union expanded to 50 states to determine if there is validity to the claim that "our politicians just don't work together like they used to." By using a two-network approach, we find compelling evidence that Congress has indeed become more polarized over time. This change has proceeded regardless of which political party controls Washington or if there is split control. Our results do not bode well for our democracy, but they do support our initial hypothesis that Congress has indeed become more polarized.

1 Introduction

Many people believe that politics has become more polarized in recent years. If this is true, it may have predictive power over the future of our nation. It can perhaps help predict how Congress will vote on controversial bills, as well as how much bipartisan action will be seen from a Congress of the future. Looking at trends can help us analyze if polarization is simply increasing, or if it increases and decreases in waves, or instead is only an imagined trend.

Our goal is to examine the evolution of polarization and collaboration over time in Congress. One way to approach this is to isolate communities in Congress and see how often these groups work together. As previous works have shown, using graph metrics to determine communities often works better at determining political divisions than simply analyzing the party divide [3]. We will first look at collaboration between communities and see how polarized each individual Congress is, and then analyze these results over time.

We define collaboration as different individuals or communities voting the same way on bills. Regardless of the motivations behind the votes, both agents are working towards the same end, and are thus collaborating in some form. Through this analysis we hope to see how polarization and collaboration have evolved over time, from the first Congress with 50 states (after Hawaii and Alaska officially became states) to the most recent Congress.

2 Related Work

2.1 Party Polarization in Congress: A Network Science Approach (Waugh, A. et al.)

Polarization is a major issue in politics. While many see the polarization between parties, many believe the issue is much more complicated. Votes in the Senate and in Congress sometimes split across party lines, and not all politicians act rationally. One approach to this ambiguity is to ignore party lines and evaluate groups some other way. One study attempted to approach this by using the concept of modularity to partition a graph into groups. Modularity works by creating partitions where there are many ties within groups and few between different groups. The benefits to this approach is that it does not assume rationality or spatial structure. Instead, it independently identifies relevant political groups and evaluates these divisions.

The network the authors built made each legislator a node in the network with an edge to each other legislator. Each edge had a weight that signified the level of agreement between these two legislators. This network, after being partitioned, was found to be useful in predicting changes in the majority party in the next Congress. The study found that at medium rates of modularity—that is not too strong or too weak—change was most likely to occur. The theory behind this is that there is partial polarization where there is some impetus for change, along with relaxed majority cohesion. The study also found a connection between polarization and re-election. Those who were divisive had a negative impact on their re-election chances, but this could be mitigated by in-group solidarity (increased polarization).

2.2 Modularity and Projection of Bipartite Networks (Rudy Arthur)

This article dives into ways to work with bipartite graphs, as they are an under-utilized and under-researched type of network. The paper discusses in detail the ways in which formulas, algorithms, and metrics differ when talking about these specific types of graphs. A lot of typical ways one interacts with a graph does not work with bipartite graphs because formulas and equations do not take into account the rules of a bipartite graph. Specifically, this paper talks about the ways in which projection and modularity differ when discussing bipartite graphs, and how one needs to take this into account when analyzing this type of graph. Modularity is difficult to measure in bipartite graphs because the graphs are already in two communities, but these communities are not "relevant" (because the nodes in the graph relate to two different subjects, it is not important when reviewing the data to notice that they are separate).

This leaves two options on how to manage bipartite graphs: modifying unipartite algorithms to account for the bipartite structure or projecting the bipartite graph onto one of its node sets. The paper opts to explore the latter, and to explain what information is preserved and lost in the various ways of converting the network to allow for standard unipartite algorithms to be used.

2.3 Modularity and community detection in bipartite networks (Michael Barber)

This paper discusses new ways to look at modularity given the complexities of the bipartite graph. This particular model uses a null model deemed appropriate for bipartite networks, and uses this model to define a form of modularity appropriate for bipartite networks. The bipartite modularity is presented in terms of a modularity matrix B, which is used to define an algorithm for identifying modules in bipartite networks. The algorithm is based on the idea that the modules in the two parts of the network are dependent, with each part mutually being used to induce the vertices for the other part into the modules. Real world results support the algorithm's usefulness in identifying known modularity in bipartite graphs.

3 Methods

3.1 Data Collection

All data was taken from `https://voteview.com/`. We used vote data from 8 different sessions of the House of Representatives, evenly spaced between when Hawaii and Alaska gained statehood and the current day. We used every 4th Congress (each 8 years) from the 87th to the 115th. The vote data was then transformed into graph data using SNAP. Each data set was used to construct two different graphs. The first is a bipartite graph, with nodes for each bill and each member, with members and bills being separate types. An edge was formed between a member and every bill with different types for yea, nay and abstained votes. Other vote types were condensed into these three vote types, making paired yea and announced yea the same as yea, and the same true for nay.

The second graph is a fully connected weighted graph between all members in the dataset, constructed similarly to the graph in "Party Polarization in Congress: A Network Science Approach" (Waugh et al.). The weight of an edge between two given members is the number of times they voted differently on bills subtracted from the number of times they voted the same way. For example, if members A and B both voted ‘yea’ on bills 1, 2, and 3, voted differently on bills 4 and 5, and both abstained from voting on bill 6, we consider their edge weight to be 2. More formally, let V be the adjacency matrix of the bipartite graph described above, where each entry V_{ib} is a integer code corresponding to the way Congressmember i voted on bill b. Then we can construct a weighted adjacency matrix A between Congressmembers where each entry A_{ij} is defined as follows

$$A_{ij} = \sum_{b \in B} [\mathbf{1}(V_{ib} = V_{jb}) - \mathbf{1}(V_{ib} \neq V_{jb})]$$

This construction allows us to think of all members of Congress as being linked by relationships somewhere along the continuum from "diametrically opposed" to "in perfect agreement." Thinking of the network in this way helps set up the problem of community detection by allowing us to skirt the question of "How strong must a tie be between two members be for them to be in the same community?" Instead of setting a relatively arbitrary cutoff for discretizing tie strength between in-group members and out-group members, this construction enables us to utilize the full continuous range of tie strengths and community strengths.

3.2 Algorithms

3.2.1 Bipartite Recursively Induced Models (BRIM)

To work with the bipartite graph, we used the BRIM algorithm discussed in [2]. This algorithm works recursively to identify bipartite modules. From the bipartite adjacency matrix A and the probability matrix P created from the bipartite graph, we can create a modularity matrix B (where $B = A - P$) and an assignment matrix S, where if a node is in a given cluster it is equal to 1, otherwise it is equal to zero.

We then partition S. Let R be a matrix that contains the assignment for only the legislator nodes, and T be the matrix that contains the bill assignments. The modularity Q can then be given by

$$Q = \frac{1}{m} Tr R^{T} T$$

, where $\widetilde{B}$ is the upper right corner of the B matrix since we only need the edges between legislators and bills, and r is the rank of $\widetilde{B}$.

The best cluster arrangement is then found through an iterative system. Nodes are first randomly assigned to clusters to build S, from which R and T can be extracted. Modularity should be maximized as per the equation given above. Then, the assignment of bills is fixed, fixing T. Then choose the R that will maximize the modularity. Afterward, the legislator assignments are fixed, so now we can choose the T that will maximize modularity. This process continues until modularity ceases to increase.

At the end of the iterations, S will contain the cluster assignment with the best modularity. Like other clustering algorithms, it is possible to become stuck in local maxima. Then, the optimal c is found by using a binomial search. The number of clusters found was then analyzed to see how the number of clusters changes over time.

3.2.2 Clustering Coefficient in Signed Weighted Graphs

For the weighted graph, we began by computing summary distributions over edge weights and local clustering coefficients for each of the 8 Congresses we studied. A degree distribution would not have been a helpful summary in this case because the graphs are fully connected. The more interesting and important part of computations on this network, though, was the clustering coefficient. Calculating the clustering coefficient was made more difficult not only by the fact that our graph has weighted edges but also by the fact that our graph has *negatively* weighted edges. This challenge led us to search for a formulation of the clustering coefficient that was appropriate for our graph, which we found in [6].

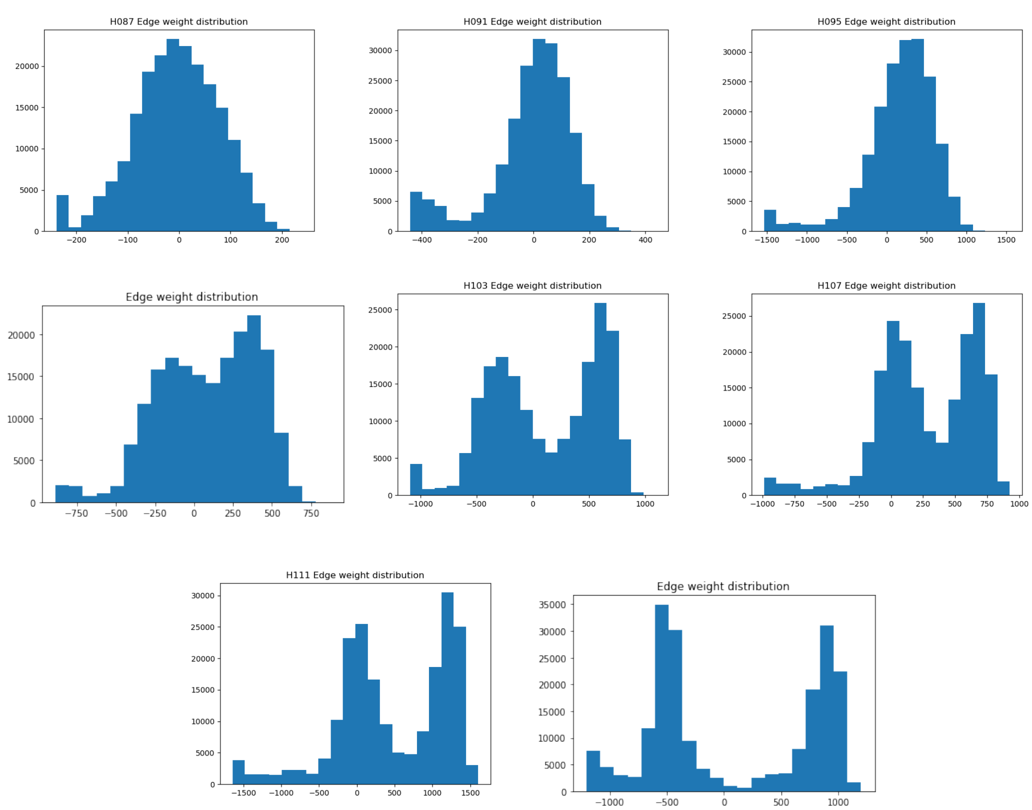
$$C_i = \frac{2}{k_i(k_i - 1)} \sum_{j,k} (w_{ij} w_{ik} w_{jk})^{\frac{1}{3}}$$

This formula just takes the normalized geometric mean edge weight for every triangle in a node's neighborhood. This gives the desired outcome of stronger ties within triangles causing higher clustering coefficients and negative ties leading to lower coefficients.

4 Results and Discussion

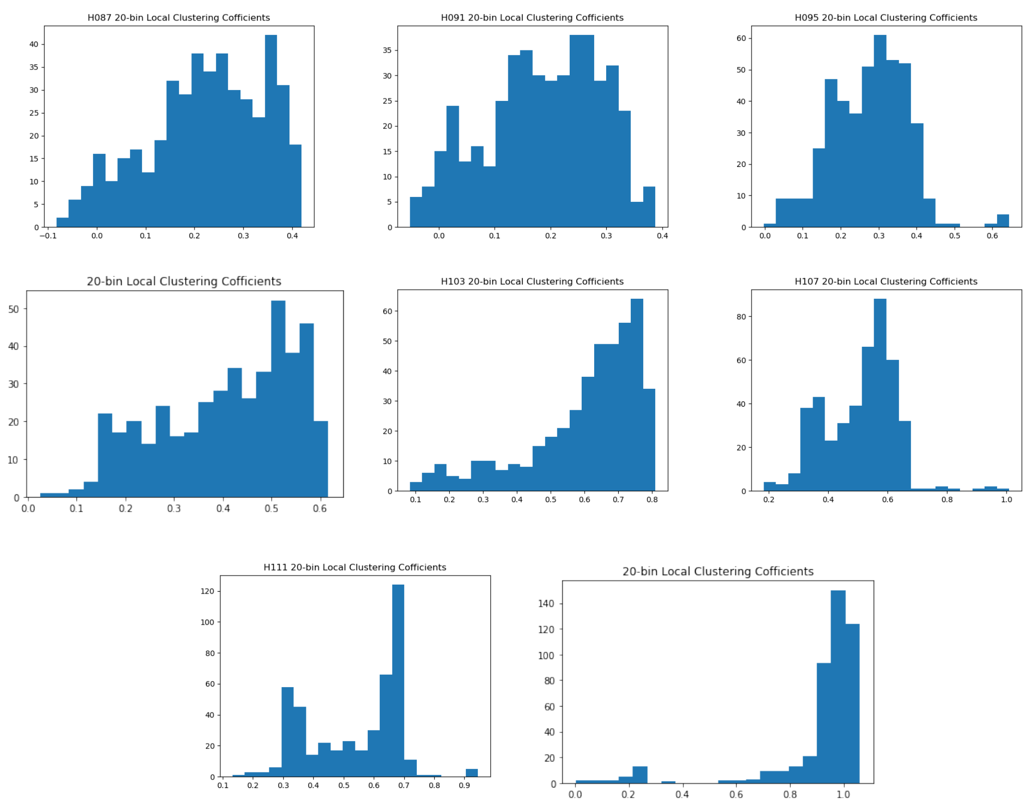
We obtained interesting results that seem to coincide nicely with our hypothesis that Congress has generally trended toward becoming more polarized over time. The first sign of support for this hypothesis comes from the weighted fully-connected graph. The edge weight distributions shifted remarkably between the 87th and 115th Houses of Representatives (see plots below). While the initial edge weights were roughly normally distributed around 0 (indicating a plurality of members working together as often as they worked against each other), the distribution shifted to a strongly bimodal distribution over the course of the eight terms we researched. This concentration around large positive and large negative edge weights shows that Congressmembers separated into clearer groups in which each member almost always voted with the same group and against the other.

Edge Weight Distributions Over Time



The next sign that Congress has become more polarized is that the distribution of local clustering coefficients has shifted towards a more clustered graph over time. This finding also comes from our weighted, fully-connected graph. As we can see in the figures below, the first three sessions we analyzed featured local clustering coefficients in roughly the 0-0.4 range, with a few outliers. Then, in the fourth and fifth sessions, the high end of the range suddenly jumps to 0.6 and then 0.8, with the high end also being the mode. The sixth and seventh sessions regress a bit in terms of clustering, but then the final one becomes incredibly clustered with a mode of 0.95-1.0. This increased clustering of the graph suggests that members are becoming more aligned with their own group and more misaligned with the other group, pointing again to increased polarization.

Local Clustering Coefficient Distributions Over Time



The BRIM method, however, was less successful. The goal of using the BRIM algorithm was to determine the optimal number of clusters for each tested term, and the modularity using that number of clusters based on the bipartite graph. However, BRIM does not support weighted or directed graphs, as mentioned in [2]: "A pitfall of BRIM, as acknowledged by Barber, is that it only handles unweighted and undirected bipartite networks." Our dataset is not weighted, but it acts like it is due to the voting options (yea, nay, and abstain must all be treated differently). To amend this, we broke the data up and performed BRIM on each dataset twice - once for the yea data and another time for the nay data. This, however, did not produce ideal results, because each run of the algorithm was missing important data, which led to vastly different estimates for the ideal number of clusters using the same congressional term data. For example, using only yea data, BRIM estimated that the ideal number of clusters of congresspeople in the 87th congressional term was 107, while nay data predicted the ideal number of clusters was 5. These both seem to miss in the wrong direction, but there's no way to determine a better prediction for ideal clustering coefficient using this method.

Conclusion

We are encouraged by the fact that our results support our initial hypothesis that Congress has become more polarized, even though we are dismayed at the current trend. We found the papers we discussed in the "Related Work" section extremely helpful in giving us novel techniques to analyze the networks we chose. At the same time, we acknowledge that having to implement those techniques from scratch was more challenging than using an off-the-shelf library like `snap.py` and perhaps took time away from finding even more angles from which to approach the question. Overall though, we feel fortunate that we were able to focus on algorithms and findings rather than on data cleaning and preparation, especially in the context of a single-quarter project.

References

[1] Arthur, Rudy. “Modularity and Projection of Bipartite Networks.” arXiv.org, August 7, 2019. `https://arxiv.org/abs/1908.02520`.

[2] Barber, Michael J. "Modularity and community detection in bipartite networks." arXiv.org, February 1, 2008. `https://arxiv.org/pdf/0707.1616.pdf`.

[3] Downs, Anthony. 1957a. An Economic Theory of Democracy. New York, NY: Harper Row. —. 1957b. “An Economic Theory of Political Action in a Democracy.” The Journal of Political Economy 65(2):135–150

[4] Hązła, Jan, Yan Jin, Elchanan Mossel, and Govind Ramnarayan. “A Geometric Model of Opinion Polarization.” arXiv.org, October 11, 2019. `https://arxiv.org/abs/1910.05274`.

[5] Lewis, Jeffrey B., Keith Poole, Howard Rosenthal, Adam Boche, Aaron Rudkin, and Luke Sonnet (2019). Voteview: Congressional Roll-Call Votes Database. `https://voteview.com/`.

[6] Onnela, J.P., et al. “Intensity and Coherence of Motifs in Weighted Complex Networks.” ArXiv.org, 7 Feb. 2005, arxiv.org/abs/cond-mat/0408629.