

# Evaluation of Political Party Cohesion Using Exponential Random Graph Modeling

*Shambavi Sadayappan*  
Whiting School of Engineering  
Johns Hopkins University  
Baltimore, MD  
ssadaya1@jhu.edu

*Ian McCulloh*  
Applied Physics Lab  
Johns Hopkins University  
Laurel, MD, USA  
ian.mcculloh@jhuapl.edu

*John Piorkowski*  
Applied Physics Lab  
Johns Hopkins University  
Laurel, MD, USA  
jpiorko2@jhu.edu

**Abstract—** The United States is becoming increasingly politically divided. In addition to polarization between the two major political parties, there is also divisiveness in intra-party dynamics. In this paper, we attempt to understand these intra-party divisions by using an exponential random graph model (ERGM) to compute a political cohesion metric to quantify the strength within the party at a given point in time. The analysis is applied to the 105th through 113th congressional sessions of the House of Representatives. We find that the Republican party not only generally exhibits stronger intra-party cohesion, but when voting patterns are broken out by topic, the party has a higher and more consistent cohesion factor compared to the Democratic Party. **Keywords—** network analysis, congressional voting patterns, exponential random graph modeling

## I. INTRODUCTION

As society moves towards hyper connectivity, whether through smart phones, social media, or a 24 hour news cycle, one aspect of life that seemingly is getting more and more disconnected is politics. In particular, the divide between the two major political parties in the United States continues to expand. Turn on the television, and you are greeted with a panel of 'experts' from each side trying to yell over one another. Go on twitter, and it will only take you a couple seconds to find a tweet pitting the conservative right against the liberal left, in a virtual shouting match with no respect for opposing viewpoints. The divisiveness has jumped from simply a nagging buzz, to a life altering force. After the highly politicized and emotional election of 2016, 16% of people have ended a close relationship due to political differences [11]. More than 77% of voters feel Americans are less tolerant of each other's political preferences than before [1]. This uptick in the political divide can clearly be seen across the two major parties. However, through the primary season of the 2016 Election, it was apparent that there were also quite significant intra-party differences of opinion. These divides manifested in political slogans such as 'Bernie or Bust' and 'Never Trump'. After the election, the Democratic party was blamed for having a messaging problem. Clinton was said to have focused too much on the negative characteristics of her opponent as a unifying message, instead of political stance to form the unifying message [9]. This analysis of the election is hypothesizing that the democratic party lacked party cohesion, which significantly impacted its performance during the

election. This underlying issue of party cohesion is the focus of this paper. This work aims to answer the following questions:

- RQ1: How can party cohesion be modeled?
- RQ2: Does party cohesion fluctuate with time?
- RQ3: Can the core values of a party be identified by observing party cohesion metrics on specific topics over time?

These questions are explored using statistical network modeling using voting data from nine sessions of the House of Representatives: from the 105th Congress through the 113th Congress.

## II. BACKGROUND

### A. Related Work

The most prominent use of statistics to analyze roll call data is through the use of the NOMINATE system [6]. This groundbreaking system used multi-dimensional scaling techniques on political data in order to assign scores to various politicians. This data has been used to show the increasing polarization in our politics. Researchers have recently started applying social network analysis to seek insights on the underlying and driving forces within Congress. The focus has been more heavily placed on looking at committee and caucus member relationships. Porter et. al [7] used roll call data and SVD (Singular Value Decomposition) analysis to identify correlations between political position and committee membership. Victor and Ringe [10] modeled the House of Representatives by creating a caucus system which connected members who were in the same caucus. Their conclusions contradict the conventional wisdom regarding the function of caucuses as a platform to provide the minority a more prominent role in legislation. Researchers have used co-sponsorships to create networks between congressional members. Fowler [3] used a co-sponsorship network and centrality and connectedness measurements to identify the most influential legislators on the hill. Our work attempts to expand the scope of this intersection of network and political science, by using ERGM analysis to look into intra-party cohesion.

## B. ERGM

Exponential-family random graph modeling (ERGM) [5] is a statistical modeling approach used to characterize social networks. The observed network in question is one of many possible networks. ERGM attempts to determine what parameters or structural characteristics of the network help shape it. These influential factors include homophily, transitivity, reciprocity, amongst others. These parameters are used to create the model. A probability distribution is formed to represent all possible networks and their respective probabilities of occurring under the model. This distribution is represented in ERGM as:

$$P(Y = y|\theta) = \frac{\exp(\theta^T s(y))}{c(\theta)} \quad (1)$$

where  $Y$  is a network,  $s(y)$  is a vector of network parameters,  $\theta$  is a vector of coefficients, and  $c(\theta)$  is a normalizing constant. The log odds probability of a tie occurring in the network is represented as:

$$\text{logit}(Y_{ij} = 1|Y^c) = (\theta)' \delta s(y)_{ij} \quad (2)$$

where  $Y_{ij}$  represents an actor pair in the network  $Y$ ,  $Y^c$  is the rest of the network, and  $\delta s(y)_{ij}$  is the change in the vector of network statistics caused by setting  $Y_{ij}$  to 1[4][8].

## III. DATASET

### A. Data Collection

This work leverages the large amount of congressional vote data available through the VoteView project [2]. Congressional member and roll call data was downloaded for the House of Representatives for the 105th through 113th congressional sessions.

### B. Network Creation

From the roll call data, an adjacency matrix was created for each congressional session. This was accomplished by creating a matrix of dimensions  $n \times n$ , where  $n$  represents the total number of house members in that specified congress. The value  $A_{ij}$  quantifies the relationship between house member  $i$  and house member  $j$ . This value was calculated by comparing the voting records of the two members, and counting the number of times they voted the same. The same approach was also used to create adjacency matrices for subset of votes which fell under specific categories. The congressional roll call data provided by VoteView has pre-classified each vote to fall under one of the following 6 Clausen categories:

1. Government Management
2. Social Welfare
3. Agriculture
4. Foreign and Domestic Policy
5. Civil Liberties
6. Miscellaneous Policy

An adjacency matrix was created to represent similarity of votes for each of the Clausen categories. Due to the very small number of votes in each session for agriculture, this category was not included in the analysis. For each session of congress, six adjacency matrices were constructed, one representing all votes for the session, and the other five representing votes restricted by topic.

## IV. PARTY COHESION ANALYSIS

With regards to RQ1, ERGM was used to calculate a party cohesion metric. Using ERGM, a model was created for each congressional session vote matrix, which accounted for edges along with a nodematch term for party membership. The result represents the log odds of an edge forming between two members of the same party. The corresponding probability of formation of this edge can be calculating through the inverse logit of the coefficients in the resulting model:

$$P = \frac{\exp(\exp(\text{coeff}))}{1 + \exp(\text{coeff})} \quad (3)$$

This corresponding probability will be considered the main party cohesion metric and is analyzed in this paper. The initial raw adjacency matrix represents a fully connected network, since generally every house member can be expected to vote identically as every other member on at least one vote. Therefore, running an ERGM analysis on the original matrix would not provide any insight. Due to the binary nature of the networks in ERGM analysis, a weighted network must first be converted to an unweighted (binary) network before ERGM can be applied. This is usually done by setting a minimal weight threshold for an edge in the network to be retained.

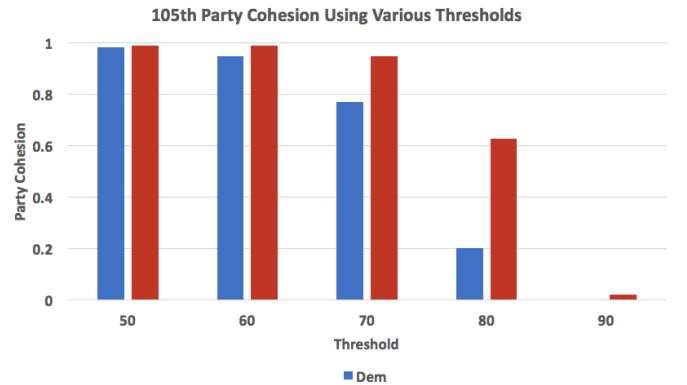


Fig. 1: Probability of a network tie in the 105<sup>th</sup> Congressional Session, tested with different thresholds

Thus, weak links between house members get removed. For the congressional voting affinity network, the threshold represents a minimum percentage of votes that two members must have in common in order to include an edge. A set of thresholds were tested on two congressional sessions: 105th congress, where the Republicans were in the majority, and the 110th congress, where the Democrats were in the majority. Figure 1 represents the output from ERGM analysis of networks derived from using different thresholds, for the 105th Congress. Figure 2 shows similar results for the 110th Congress. As can be seen in these figures, the party in power

has a higher probability of edge formation, irrespective of the chosen threshold to construct the binary graph. Moving forward in our analysis, the threshold was set to 80%, i.e., in order for two house members to be connected, they must have voted identically on at least 80% of the votes.

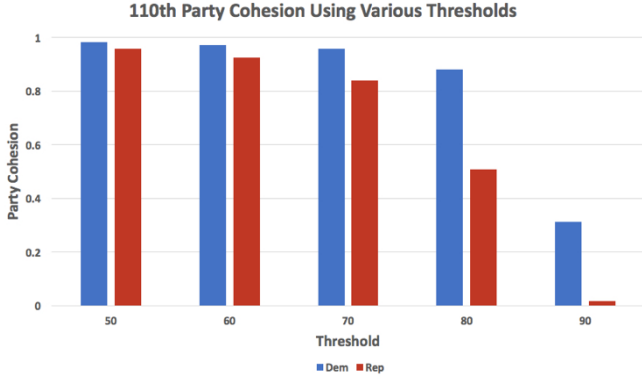


Fig. 2: Probability of a network tie in the 110th Congressional Session, tested with different thresholds

## V. PARTY COHESION OVER TIME

In order to answer RQ2, we compare the party cohesion metric for an individual party across the span of 9 congresses. Using the threshold of 80%, the overall intra-party cohesion within the Democratic Party and Republican Party was plotted against the sessions of congresses. Figure 3 represents the results of this comparison. As seen in the graph, the two parties fluctuate between who has a higher cohesion. The red background in the graph represents the period over which the Republican party held the House majority, while a blue background indicates a period of Democratic House majority. In the past 9 congresses, the two instances where one party overtook the other party with regards to political cohesion coincided with the stronger party taking control of the House of Representatives.

## VI. TOPIC BASED COHESION ANALYSIS

To answer RQ3, we must analyze the party cohesion for specific topics. This is done by creating ERGM models for each individual topic vote matrix for each party, across the range of congressional sessions. Figure 4 shows the results of this analysis for the Democratic Party and Figure 5 shows the results for the Republican Party. The overall fluctuation in party cohesion for the Democratic party is generally larger than for the Republican party. This can be seen more clearly in

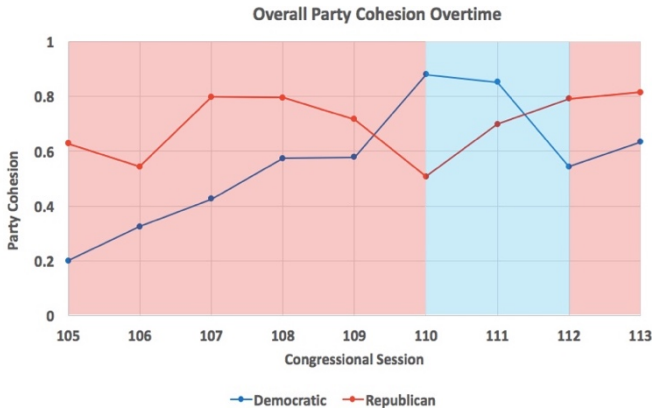


Fig. 5: Party cohesion over time

Figure 6, which illustrates the fluctuation range for each topic. The bars represent the range, while the isolated line fragments represent the average party cohesion probability for that specific topic. The Democratic party has a larger range of fluctuation for every topic except Social Welfare. They also have, on average, a lower topic cohesion than the Republicans.

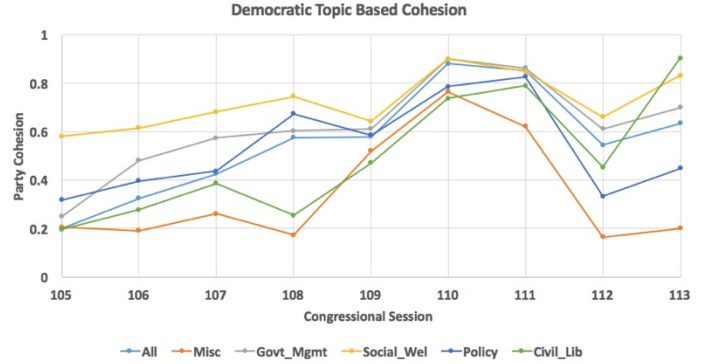


Fig. 3: Topic-based cohesion for the Democratic Party

## VII. DISCUSSION AND CONCLUSION

### A. Key Findings

In this paper, we set out to utilize ERGM to get a better understanding of the political cohesion of the two major political parties in the United States, and attempt to identify any longitudinal trends with regards to overall cohesion, as well as topic-based cohesion. We were able to use results of the ERGM modeling to calculate the overall probability of two members of a party voting together a certain number of times in a congressional session. The probability that two members of a party voted together at least 80% of the time became the metric by which we measured intra-party cohesion. We found that for the congressional sessions 105 through 113, the party with the higher cohesion fluctuated. Each time a party took over as the more cohesive party, they became the majority in the House of Representatives. When analyzing topic-specific party cohesion, we found that, in general, the Democratic party is less cohesive than the Republican party. This was found not only in average cohesion, but also in range of fluctuation. This indicates that the Democratic party does not have a united voice when it comes to policy issues, government

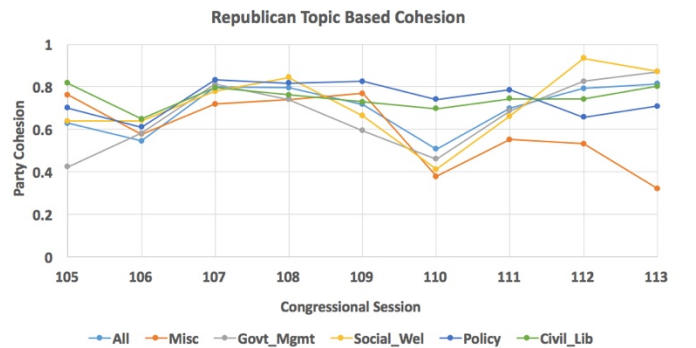


Fig. 4: Topic-based cohesion for the Republican Party

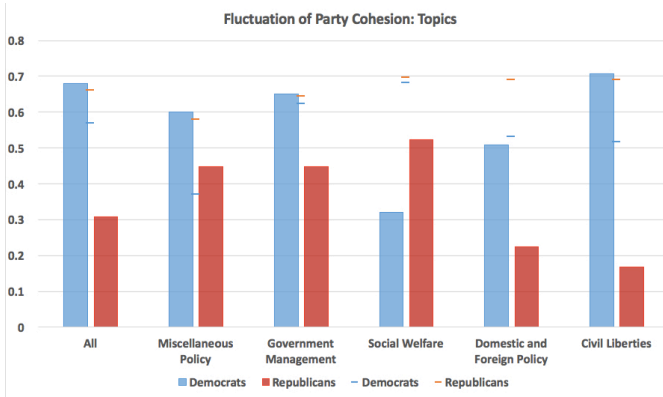


Fig. 6: Fluctuation of party cohesion over various topics

management, or civil liberties. The topic which seems to be the bedrock, fundamental core of the Democratic Party is Social Welfare. On the other hand, Republicans seem to have a highly cohesive party across the spectrum of topics. This indicates that the Republican party in the House of Representatives has a more cohesive and unifying platform compared to the Democratic party.

The results from the topic-specific cohesion analysis support the moral foundations hypothesis laid out by Graham et al. [12]. The moral foundations hypothesis posits that there are five core moral foundations, and liberals and conservatives base their morality on different configurations of these foundations. Liberals are focused two core morals: harm/care and fairness. Conservatives, however, have a moral foundation that more equally weights all five different core morals. The results from the topic-specific party cohesion models pointed towards a Democratic party that comes together over Social Welfare and a Republican party that is quite united over a variety of topics, ranging from Social Welfare to Foreign Policy.

### B. Limitations

The political landscape is a complicated system, in which the House of Representatives is only one piece. This paper only focused on a small subset of time and people involved in United States politics.

## VIII. FUTURE WORK

This paper only touches the surface in analyzing intra-party cohesion. Future work can expand the timeline, players

involved, and attributes analyzed. One planned direction of study is the analysis of only contested votes, where the majority view of both parties was in opposition. Such an analysis might reveal sharper distinctions than analysis of the entire set of votes.

## IX. ACKNOWLEDGEMENTS

This work was supported by the Office of Naval Research, Grant No. N00014-17-1-2981/127025

## REFERENCES

- [1] "40 percent say election has hurt a close relationship.", Rasmussen Reports, Feb 2, 2017. [Online]. Available: [http://www.rasmussenreports.com/public\\_content/politics/elections/election\\_2016/40\\_say\\_election\\_has\\_hurt\\_a\\_close\\_relationship](http://www.rasmussenreports.com/public_content/politics/elections/election_2016/40_say_election_has_hurt_a_close_relationship)
- [2] Voteview database. <https://voteview.polisci.ucla.edu/>.
- [3] James H Fowler. Legislative cosponsorship networks in the us house and senate. *Social Networks*, 28(4):454–465, 2006.
- [4] David R Hunter, Mark S Handcock, Carter T Butts, Steven M Goodreau, and Martina Morris. ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of statistical software*, 24(3):nihpa54860, 2008.
- [5] Dean Lusher, Johan Koskinen, and Garry Robins. Exponential random graph models for social networks: Theory, methods, and applications. Cambridge University Press, 2012.
- [6] Keith T Poole and Howard Rosenthal. A spatial model for legislative roll call analysis. *American Journal of Political Science*, pages 357–384, 1985.
- [7] Mason A. Porter, Peter J. Mucha, M. E. J. Newman, and Casey M. Warmbrand. A network analysis of committees in the u.s. house of representatives. 102(20):7057–7062, 2005.
- [8] Garry Robins, Pip Pattison, Yuval Kalish, and Dean Lusher. An introduction to exponential random graph ( $p^*$ ) models for social networks. *Social networks*, 29(2):173–191, 2007.
- [9] Kristin Tate. With no emerging leaders, no clear message, Democrats flounder. <http://thehill.com/blogs/pundits-blog/national-party-news/325444-party-of-nancy-pelosi-with-no-emerging-leaders-no>.
- [10] Jennifer Nicoll Victor and Nils Ringe. The social utility of informal institutions: Caucuses as networks in the 110th us house of representatives. *American Politics Research*, 37(5):742–766, 2009.
- [11] John Whitesides. From disputes to a breakup: Wounds still raw after U.S. election. <http://www.reuters.com/article/us-usa-trump-relationships-insight-idUSKBN15M13L>.
- [12] Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of personality and social psychology*, 96(5), 1029.