UTRECHT UNIVERSITY

Department of Information and Computing Science

Applied Data Science Master Thesis

Examining the Occurrence of Ideological and Affective Polarization within the US Congress: a Structural Balance Theory Approach

First Examiner:
Javier Garcia Bernardo

Candidate:
Esther Stehouwer

Second Examiner:
Elena Candellone

In Cooperation with:
Dion Custers & Lars Teunissen

June 28, 2024

Abstract

Background: Political polarization is the divergence of political behaviour away from the middle and towards the ideological extreme. Political polarization can move towards a pernicious or beneficial state. Ideological and affective polarization, i.e., different forms of political polarization, explain how political polarization can move towards either a beneficial or pernicious state. More knowledge on what state a country is moving to can increase our knowledge about the consequences of it. These include increase of democratization in case of a beneficial state and an increase of ideological gap and an "us" versus "them" society in case of pernicious state.

Objectives: Currently, research predominantly quantifies ideological polarization by calculating DW-NOMINATE scores, which is done by placing individuals on a scale from conservative to liberal. However, DW-NOMINATE cannot determine affective polarization. Therefore, the objective of this study was to consider political polarization as a network science problem to quantify both ideological and affective polarization. This is done by applying Structural Balance Theory (SBT), that examines the stability and consistency of relationships in social networks.

Methods: Stochastic Degree Sequence Model (SDSM) is applied to create networks. With an unsigned SDSM, weak polarization, being an indicator of ideological polarization, can be determined. Weak polarization is quantified with modularity scores. With a signed SDSM, strong polarization, being an indicator of affective polarization, can be determined. Strong polarization is quantified with Triangle Indexes.

Results: Modularity scores, and thus weak polarization, increased significantly monotonically over time for the House of Representatives. There was no significant trend for the Senate. The Triangle Index, and thus strong polarization, did not change significantly over time for the House of Representatives. However, there was a significant decrease of +++ triangles and significant increase of +-- triangles. Additionally, there was a significant increase of +++ triangles within party and significant increase of +-- triangles between parties in the House of Representatives. Therefore, the results regarding strong polarization, and thus of affective polarization, in the House of Representatives remain inconsistent. There are no results of strong polarization, and thus of affective polarization, in the Senate.

Conclusion: The main finding of this master thesis is that there is an indication for the existence and increase of ideological polarization, but not for affective polarization. However, since these are only indications, more research is needed to further develop a working method to examine occurrence of ideological and affective polarization.

Keywords: political polarization, ideological polarization, affective polarization, weak polarization, strong polarization, Structural Balance Theory, Stochastic Degree Sequence Model.

Table of Contents

1 Introduction

Polarization can be defined as the division of a community into two clearly contrasting groups (e.g., the rich and poor; nationalists and globalists; and traditionalists and modernists) (1). Polarization exists also within the political realm and is defined as political polarization, which entails the divergence of political behaviour away from the middle and towards the ideological extreme (2). The implications of political polarization are crucial to understand, because it can provide insights in the (dis)functioning of governments and democratic institutions. These include both beneficial consequences as well as detrimental outcomes.

In its most extreme forms political polarization can have destructive consequences, which is referred to as pernicious polarization (3). In a perniciously polarized society, there is a deep ideological gap, in which differences between groups have become too immense to reach a compromise (3). Individuals are utterly hostile towards each other and society has become an "us" versus "them" society (3). This is a result of extreme feelings of distrust and fear towards each other (3). Additionally, individuals tend to dehumanise and stereotype members of the opposing community (3). An "us" versus "them" society is further reinforced by the existence of echo chambers (4). This is an environment in which individuals only encounter information that reinforce their views and beliefs (4). For instance, echo chambers are formed because algorithms prioritize content similar to what users have previously viewed (4). Heltzel and Laurin stated that pernicious polarization can lead to social fragmentation, increase in political violence, economic stagnation, and distrust in the government (3). The latter can already be noticed: the public trust in government has decreased from 73% in 1958 to a striking 16% in 2023 (5).

On the contrary, political polarization can also have benefits. Polarization, in its simple forms, can contribute to democratization in multiple ways. Firstly, McCoy and Somer state that polarization can increase political awareness and participation (6). They explain that in a more polarized country, it is easier to make well-considered decisions on who to support, since the different ideologies create clear distinctions, thus increasing political awareness (6). Additionally, because of the more distinct ideologies, people tend to be more involved with the party that follows their ideology, thus increasing political participation (6). Weber et al., explain that a more polarized country can also lead to an increase in political competition (7). Consequently, this leads to innovation in policy and governance, since people want the best for their party (7).

It is essential to highlight the distinction between beneficial and pernicious polarization. In countries where polarization is perceived as beneficial, individuals tend to prioritize within group relationships and are less concerned with out-group relations. The focus on internal relations stimulates individuals to strive for personal improvement, fostering a competitive yet constructive environment. On the contrary, in countries where polarization is perceived as pernicious, individuals are fixated on the out-group relations. They try to improve by also depreciating the other group, leading to a hostile and adversarial environment.

The above shows the relevance on studying the occurrence of political polarization within a country, since it can have both beneficial as well as destructing consequences. Gaining more insights into the type of polarization a country is moving to (beneficial or pernicious sight) can help us better understand political polarization over time within a country. Then, the question arises about what kinds of political polarization contribute to the occurrence of

polarization in both positive and negative manifestations. Ideological and affective polarization have shown to be the different types that can be related to political polarization in its beneficial and pernicious form respectively (8). In the literature review below, this will be more elaborated upon. Research has already been done extensively on the existence of ideological polarization in a country (9). However, to the best of my knowledge, little research has been done in which both ideological and affective polarization are quantified. Therefore, this thesis aims to expand on researching both ideological and affective polarization. Because of the abundance of data availability on the legislators within the United States (US) Congress, this thesis focuses on political polarization within this context (10,11). In consideration of the above, it is worth investigating the following research questions: how does political polarization manifest in the United States Congress from the 93rd to the 117th sessions, and to what extent is there evidence of ideological and affective polarization in the interactions between Democratic and Republican legislators?

2 Literature Review

2.1 Ideological and Affective Polarization

Countries can be politically polarized in multiple ways. Political scientists distinguish between mass polarization and elite polarization. Mass polarization refers to a society with a strong division in political parties, in which members of the party oftentimes have very opposing point of views (12). This can lead to polarisation within the electorate, since citizens create their view on the world through the lenses of the political party and have little to no sympathy for the opposing party (13). Hence, the effectivity of political processes can also be diminished (13). Elite polarization can be defined as the sharp division between the in-government party and the opposing party (14). The parties are internally very cohesive, meaning adequate cooperation between the members to operate as a unity through programmatic behaviour, consisting of a consistent set of ideas and ideologies they strive for (14). Hence, the political party in power drifts further away from the conflicting party in terms of their policy, views, and ideology (14). Zingher explains that elite polarization can result in amplification of mass polarization (15). Namely, strongly polarized elites can further polarize the members of their party by showing their hostility towards the opposing party onto their members through speeches, policy, or other public facilities (15). Additionally, research showed that elite polarization can lead to the introduction of more extreme policy proposals and less cooperation between the parties, which leads to stagnation in handling legislation and policy (14).

The above mainly focusses on elite and mass polarization as being forms of ideological polarization, which can be defined as members of the electorate (in case of elite polarization) or the party (in case of mass polarization) have different views on ideological issues (e.g., abortion or death penalty) (8). Under this lens, members of the party or electorate are not necessarily extreme, but rather consistent in their positions (8). Thus, members of opposing parties or electorate are divided simply because of their differences in ideology. Contrastingly, polarization can also arise because members of the party effectively dislike members of the contrasting party and feel hostility towards them, a phenomenon referred to as affective polarization (14). It is important to note that feelings and emotion distinguish ideological polarization from affective polarization: the latter is determined by the occurrence of in-group favouritism and out-group derogation (14).

2.2 DW-NOMINATE

Extensive research already has been done on measuring political polarization within the US congress. A well-known application of measuring ideological polarization is by calculating DW-NOMINATE scores. Introduced by Poole and Rosenthal, this is a method in which the roll-call votes are used to place the legislators on an ideological scale (16). These scores can be two-dimensional, but generally the DW-NOMINATE scores are placed on a one-dimensional scale with the scores varying between -1 and +1, referring to extremely liberal and conservative respectively (17). This method provides a quantitative way of measuring ideological positions of legislators (16). Ever since, it has been a widely accepted way of measuring ideological positions for multiple reasons. Firstly, the method allows to measure into relative positions of both individual as well as major parties (e.g., democrats and republicans). Additionally, the two-dimensionality of the method allows for comparison for multiple topics, which can provide

meaningful information on whether polarization is consistent across multiple subjects (e.g., economics, defence).

2.3 Structural Balance Theory

On the contrary, a considerable implication of DW-NOMINATE is that this method can only provide values related to the ideology of a legislator, based on their voting behaviour. Thus, DW-NOMINATE scores only allow for quantification of ideological polarization. This gave rise to the question how to measure other forms of political polarization, such as affective polarization. Neal answered this question by considering the existence of political polarization as a network-science problem (9). In his paper, Neal suggests that one way to quantify affective polarization is by examining the existence of positive relations between members within party and negative relations between members of opposing parties, which he refers to as strong polarization (9). Neal considered strong polarization as a specific phenomenon arising within Social Identity Theory (SDI). According to SDI, which was proposed firstly by Tajfel and Turner in 1979, the identity of individuals is partly based on their memberships of different communities (18). Additionally, individuals strive for a positive self-image that can be created by putting emphasis on the within-community ideals and beliefs, and by differentiating from ideals and beliefs of other communities than their own (18).

Neal in his article proposed the application of Structural Balance Theory (SBT) to visualise the existence of the in-group favouritism and out-group derogation (9). Originally developed by psychologist Fritz Heider, SBT is a theory within social psychology that examines the stability and consistency of relationships in social networks (19). It focuses on relationships between three individuals that form a triad of positive (+) and/or negative (-) relationships. A triad can be either balanced or unbalanced. Balanced triads are those triads in which all three individuals have a positive relationship with each other (+++) or triads in which there is one positive relationship and two negative relationships (+--). The latter triad refers to the concept of "the enemy of my friend is also my enemy." Unbalanced triads are those triads in which all three individuals have a negative relationship with each other (---) or those triads in which there are two positive relationships and one negative (++-) (see also Figure 1).

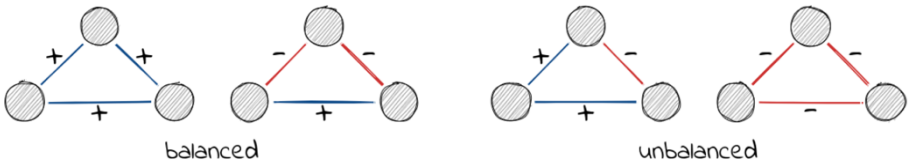


Figure 1. Triads of Structural Balance Theory

Neal suggested that the balanced triads could be used as a way to measure affective and ideological polarization depending on how one interprets the positive and negative relationships of the triads (9). Neal refers to weak and strong polarization, in which the first can be defined as the in-group favouritism *without* out-group derogation and the latter as the in-group favouritism *with* out-group derogation respectively (9). Neal suggested that the balanced triads mirror weak and strong polarization respectively. He states that +++ triangles occur within networks with great party affiliation, which he defines as weak polarization. When considering +++ triangles in this way, this could contribute to a way of measuring ideological polarization.

More interestingly, Neal declares that one way to interpret the balanced triangles is by stating that the +-- triangles occur within networks in which individuals become friends because they dislike the same person. Thus, it is exactly this triad that can reveal the existence of in-group favouritism and out-group derogation, i.e. the existence of what Neal defines as strong polarization. When considering the balanced triangles in this manner, it could contribute to a way of measuring affective polarization. However, it is important to state that interpreting the balanced triangles in this way is just one way of interpreting it. Evidently, this means that weak and strong polarization are not a one-to-one relation with ideological and affective polarization respectively. Nevertheless, it is a realistic way of measuring ideological and affective polarization. Thus, Neal proposed a method to examine weak and strong polarization, which could contribute to measuring ideological and affective polarization respectively, which goes beyond the DW-NOMINATE method. This thesis aims to examine the occurrence of both ideological as well as affective polarization, hence we will apply the method suggested by Neal.

3 Data

3.1 Data Preparation for Analysis

There are three datasets that need preparation for analysis: the roll call dataset; the members dataset; and the member roll call votes dataset. Extraction of the datasets can be done via www.voteview.com, which is an open-source website. Therefore, there are no ethical or legal considerations that needed to be considered before using the data.

The Roll Call Dataset

The initial roll call dataset is retrieved from www.voteview.com and is filtered from the 93rd until the current 118th congress, since information from before the 93rd congress is too minimal (20). However, the information contained in this filtered roll call dataset is still too little to perform proper topic modelling. Therefore, extra information on each roll call is included by adding information from the roll call dataset from www.propublica.org (21). This extra information entails long summaries for each roll call. This expands the lexicon in the roll call dataset, enabling useful performance of topic modelling. Roll calls without summaries are removed, since these must exist to execute topic modelling. This results in the roll call dataset containing only bills. Information is added (e.g., policy and subjects) to control for the coherence of the topics found. The final roll call dataset contains of 28,449 bills for the 93rd until 118th congress.

The Member Dataset

The initial member dataset is retrieved from www.voteview.com and is filtered from the 93rd until the current 118th congress. Every member of the dataset receives a new unique ID, since some members in the initial member dataset turn out to have multiple IDs. The final member dataset contains 11,580 unique rows, of which 2217 are unique persons and 2246 are unique combinations of person and party.

The Member Roll Call Votes

The initial member roll call votes dataset is retrieved from www.voteview.com and is filtered from the 93rd until the current 118th congress. The bills and the unique IDs are added within the member roll call votes dataset to enhance integration of the three datasets. The final member roll call votes dataset contains 12,343,235 member votes.

The Legislators Datasets, Bills Datasets, and Matrix Datasets

Legislators datasets and bills datasets are retrieved from filtering the right information from the member roll call votes. For the legislators datasets this includes the legislator's names, their unique ID, their party (i.e., republican, democrat, or independent), and their position (i.e., House of Representatives or Senate). A separate legislator dataset is created for congresses 93 until 117 (see below for argumentation). Thus, there are 25 separate legislators datasets. For the bills datasets this includes the bill number (i.e., bill, joint-, concurrent- or simple resolution with a unique number), the unique bill ID, the sponsor's name, the sponsor's party, and the topic it belongs to retrieved from topic modelling (i.e., defense and military, environmental and natural resources, government budget and administration, infrastructure and development, international relations and government, legislation and policy, and social services and public welfare).

Separate bills datasets are created for congresses 93 until 117. Thus, there are 25 separate bills datasets.

The matrix datasets are created to connect the bills with the (co-)sponsors of the bills. The columns represent the bill numbers and the rows the (co-)sponsors. Initially, the matrix is the null matrix. Then, for each bill a '1' is entered in the matrix where the row ((co-)sponsor) and column (bill number) intersect, if the (co-)sponsor in question is in the list of that bill. This is repeated for the congresses 93 until 117. Thus, there are 25 separate matrix datasets. Table 1, 2 and 3 and Figure 2 show an overview of the details or summary statistics of the different datasets.

Certain aspects of these three datasets warrant attention. Firstly, information on the bills and (co-)sponsorship behaviour is used for congress 93 until 117 (1973-2022). Congress 93 is chosen as a starting point, since this is the first congress of which enough electronical data is available for both chambers. Even though there is also data available for congress 118, this congress is still ongoing and therefore incomplete. Therefore, congress 117 is chosen as the ending point. Secondly, the type of bills considered in this thesis are joint resolutions and bills. Legislation can be introduced within the US Congress in four different manners: as a bill or joint resolution, which can become law after being sent to and signed by the President; or as a concurrent or simple resolution, which cannot become law since they cannot be sent to the President (9). Introduced bills that can become law provide more information on the political position and therefore, only the joint resolution and bills were considered in this thesis. Thirdly, no distinction is made between sponsor or cosponsor. The sponsor is defined as the first person that introduced the bill within the chamber (9). The cosponsors then are the legislators that support the introduction of the bill. However, the difference between sponsor and cosponsor remains subtle, since the sponsor is not necessarily also the person who created the bill, this could also have been a group of legislators who later become the cosponsors (9). Lastly, both the Senate and House of Representatives are considered, since bills and joint resolutions can be introduced within both the Senate as well as the House of Representatives. Therefore, both chambers contain information on sponsorship behaviour of legislators. However, since they are two separate chambers, they are also considered as two separate datasets.

Table 1. Summary statistics of number bills and number of legislators datasets containing minimum, median and maximum values

Dataset	***Minimum***	***Median***	***Maximum***
Bills	220 (99th congress)	489	985 (111th congress)
Legislators	491 (93rd congress)	542	582 (113th congress)

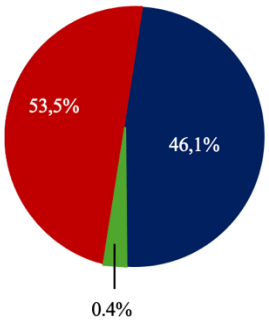


Figure 2. Pie chart of average division of republicans (red), democrats (blue) and independents in US congress.

Table 2. Different columns in bills datasets with its type and meaning

Column header	***Type***	***Meaning***
Congress	Integer. 1+	The number of the congress the member belongs to
Bill_number	String	Bill identifier
ID	String	Member identifier
Sponsor_party	String. D, R, or I	Party of sponsor. Democrat, Republican or Independent
Topic_with_coherence	String	Topic the bill belongs to

Table 3. Different columns in legislators datasets with its type and meaning

Column header	***Type***	***Meaning***
Congress	Integer. 1+	The number of the congress the member belongs to
Name	String	Name of member, surname first
ID	String	Member identifier
Party	String. D, R, or I	Party of the member. Democrat, Republican or Independent
Type	String. Rep or Sen.	The chamber in which the member served. House of Representatives or Senate

3.2 Motivation for using the Matrix, Legislators, and Bills Datasets

Previous literature use data on the (co-)sponsoring of bills because of their rich amount of information on political positions (9,22–24). Bills refer to the legislative proposals from the House of Representatives and Senate within the United States (9). Before a bill becomes law, it undergoes multiple steps: a bill is introduced within the House of Representatives or Senate by a legislator. This legislator is then called the *sponsor* of that bill. By joining the bill as a *co-sponsor*, other legislators can show their support for that bill. After the bill has been passed in the chamber, it can be sent to the other chamber, and, if passed, can lastly be sent to the president. After the president has signed the bill, it becomes law (9). However, the great majority of bills will never be voted upon (e.g., 2% in 1973-1974 and 5% in 2015-2016) (9). Contrastingly, all legislators are (co-)sponsor of a certain number of introduced bills (9). Therefore, (co-)sponsorship behaviour of legislators provides a great amount of information on their political positions and was therefore used in this thesis.

4 Methods

Research was conducted in R within the period of 22-04-2024 and 28-06-2024. R-code can be found in the GitHub using the following link:
https://github.com/estherstehouwer/masterthesis.git.
The first step that is taken is to split the matrix, legislator and bills datasets between House of Representatives and Senate, since they are two separate chambers, as mentioned above.

4.1 The Stochastic-Degree Sequence Model (SDSM)

As mentioned above, ideological and affective polarization is analysed using Neal's approach (9). Therefore, the second step that is taken is creating a network that represents the sponsorship behaviour of bills among legislators. This is done by applying the Stochastic-Degree Sequence Model (SDSM) that is found in the library *Backbone*, which was created by Neal in 2014 (9).

The main argument for applying SDSM is that it allows for researching weak as well as strong polarization, since it allows for both positive as well as negative edge weights (1 and -1). The SDSM generates a network based on how often legislators (co-)sponsor the same type of bill. It does this in four steps. First, it calculates how many of the same bills would be (co-)sponsored by two legislators at random, considering that some legislators (co-)sponsor more bills than others, and that some bills have more (co-)sponsors than other bills do. Then, the model develops a distribution of the joint bills under the null-hypothesis that cooperation between two legislators is random. Next, the actual joint number of bills that are (co-)sponsored by two legislators is compared to the expected values from the null-hypothesis. A two-tailed significance of α-level of 0.05 is applied to determine if there is a statistically significant difference between the expected and actual values. Lastly, the model creates a positive edge between two legislators when the actual values were statistically significantly more than expected, and a negative edge when the actual values were statistically significantly less than expected. The function to call SDSM has the possibility to state whether the network should be signed (both positive as well as negative relations) or non-signed (only positive relations). The network with positive relations only allows for analysis of weak polarization, whereas the network with both positive as well as negative relations allows for analysis of affective polarization. The function is first set to non-signed to allow analysis of weak polarization. After this, the function is set to signed to allow analysis of strong polarization. The *visNetwork* function in R is used to visualise the network. Nodes representing the democrats, republicans, and independents are coloured blue, red, and green respectively.

4.2 Quantifying occurrence of weak polarization

Weak polarization is defined as the existence of positive relations within groups, and the absence of relations between groups (14). Modularity is a measure that can quantify the interdependence within modules and independence between modules, and therefore, it is an appropriate measure to apply to estimate the occurrence of weak polarization (25). Thus, the third step taken is calculating the modularity using the *modularity* function in the *igraph* package (26). This modularity is calculated using Newman's formula (25):

$$Q = \frac{1}{2m}\sum_{ij}\left[A_{ij} - \frac{k_i k_j}{2m}\right]\delta(c_i, c_j),$$

in which m refers to the total number of edges within the network, k_i to the degree of node i, and A_{ij} to the weight between node i and j. For $\delta(c_i, c_j)$ it is known that

$$\delta(c_i, c_j) = \begin{cases} 1, & c_i = c_j \\ 0, & c_i \neq c_j. \end{cases}$$

Values of Q can range between -0.5 and 1. For the negative values ($Q \in [0.5\ , 0)$) it means that the division of the network is worse than would have been expected at random, meaning that there are less relations within groups than there are generated randomly. For $Q = 0$, the network's division is no better than random, suggesting there is no clear division in the network. For the positive values ($Q \in (0\ , 1]$), it means that the division of the network is higher than would have been expected at random, meaning that there are more relations within groups than there are generated randomly (26). The partition is based on party affiliation, meaning that the membership vector of the community structure is a vector containing the values 1, 2 and 3, referring to the legislator in the network being democrat, republican and independent respectively. The modularity scores are determined for each congress (93-117) and stored as a vector containing these values.

The fourth step taken is to determine the Spearman's Rank Correlation Coefficient (ρ), to determine whether there is a significant change in modularity over time. This coefficient assesses how well the relationship between change in community structure and time can be described applying a monotonic function (27). The Spearman's Rank Correlation Coefficient is calculated using the *cor.test* function with the method equal to Spearman using the vector with the modularity scores over time as input. The modularity scores over time are visualised using the *plot* function.

4.3 Quantifying occurrence of strong polarization

Strong polarization is defined as the existence of both positive relations within groups, as well as negative relations between groups (14). To determine the existence of strong polarization for each congress, the fifth step taken is to determine the degree of balance of the network. This refers to the extent to which balanced triangles occur within the network, following SBT as suggested by Neal above (9). This is done by calculating the proportion of +++ and +-- triangles, also referred to as the Triangle Index (TI) (28). The TI can be found by applying the *count_signed_triangles* function from the package *signnet,* using the **signed** network as input. This provides the count of each triangle type (+++, ++-, +--, ---) (29). The TI is then the sum of the balanced triangles (+++, +--) divided by the sum of the four types of triangles.

Since the TI is a proportion, values for TI range between 0 and 1. Additionally, in a randomly constructed network it is expected that approximately half of the network would be balanced (9). Therefore, a value greater than 0.5 indicates more balanced triangles than would be expected. The TIs are determined for each congress and stored as a vector containing these values.

The sixth step is to determine again Spearman's Rank Correlation Coefficient (ρ) to visualise if the change in balanced triangles is significant over time, this time using the vector with the TI values as input. The TI scores over time are visualised using the *plot* function. The seventh step is to visualise the proportion of +++ and +-- triangles separately. This is done by

storing the proportion of +++ and +-- over time in two different vectors and visualising these using the *plot* function.

Lastly, to analyse strong polarization, the change of positive relations (+++) within party and the negative relations (+--) are determined. This is done by filtering those legislators from the network that are part of a +++ or +-- triangle. The function *count_signed_triangles* can be used to determine the type of triangle, since it contains a column containing the number 0, 1, 2 or 3 referring to the ---, +--, ++-, and +++ triangles respectively. A data frame is created containing a column with the balanced triangles (1 and 3) and three columns containing the party of each legislator within that balanced triangle (D or R). Two vectors are created containing the proportion of within party +++ triangles and between party +-- triangles respectively. The former is calculated by counting the number of +++ triangles within the data frame for which all legislators were either democrat or all republican divided over the total number of rows in the data frame, i.e., all balanced triangles. The latter is calculated by counting the number of +-- triangles within the data frame for which two legislators are democrat and one is republican, or two legislators are republican, and one is democrat, divided over the total number of rows in the data frame.

5 Results

5.1 Statistical Results

Figure 3 shows boxplots of the number of nodes (**A** and **B**) and edges (**C** and **D**) for the network of each congress for both chambers. The number of nodes for each congress are predominantly constant in both chambers. For the House of Representatives, the minimum number of nodes is 378 (congress 93), the median is 423, and the maximum is 497 (congress 113). For the Senate, the minimum number of nodes is 85 (congress 113), the median is 117, and the maximum is 127 (congress 95) (see Table 4).

The number of edges increase over time from 1519 in congress 97 to 13247 in congress 117 in the House of Representatives. The congresses in the Senate with zero edges are the result of the SDSM being unable to create a network. For the House of Representatives, the minimum number of edges is 1519 (congress 97), the median is 4259, and the maximum is 13247 (congress 117). For the Senate, the minimum number of edges is 0 (multiple congresses), the median is 1, and the maximum is 66 (congress 117) (see Table 5).

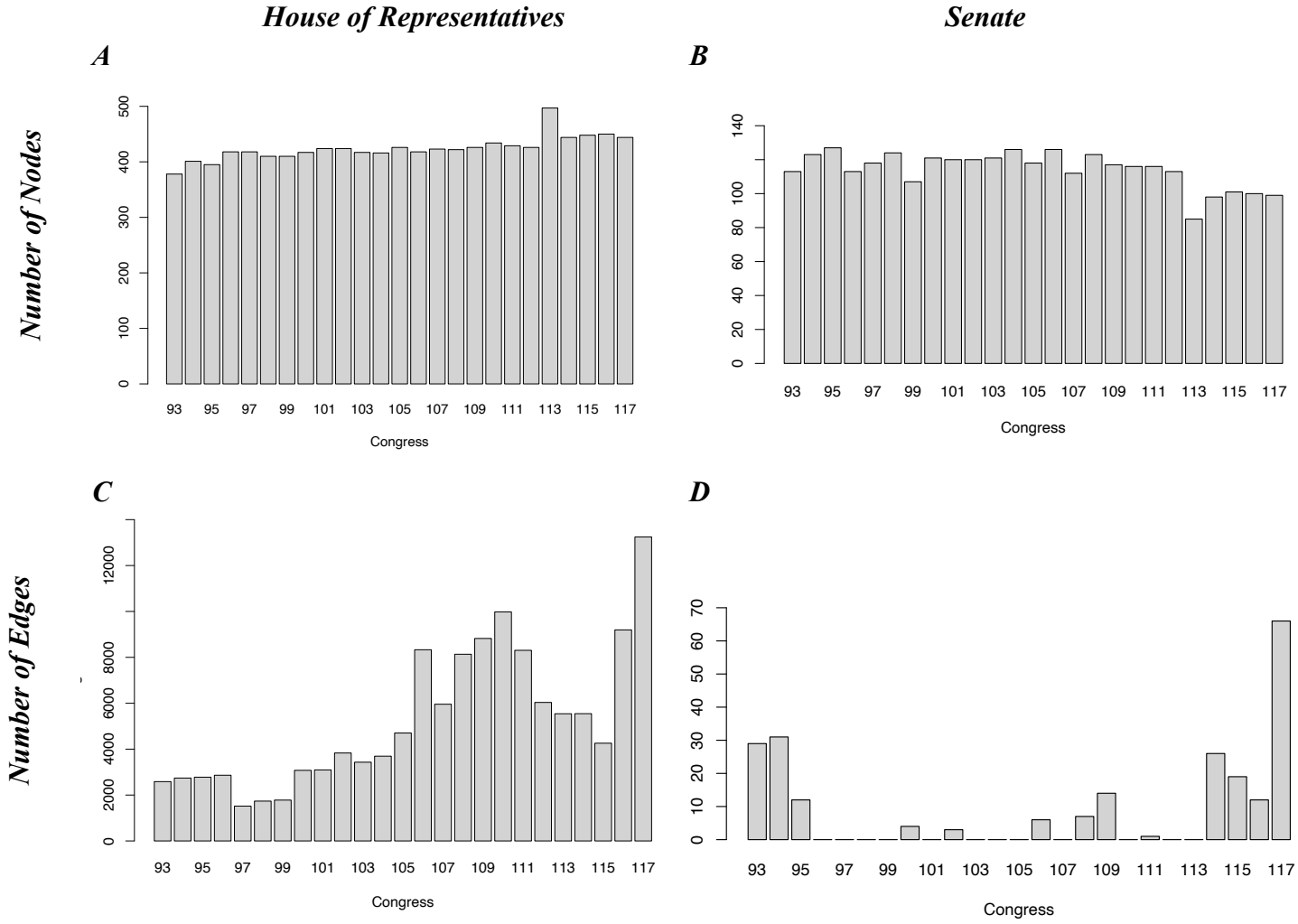


Figure 3. Boxplot of the number of nodes (**A** and **B**) and edges (**C** and **D**) for the network of each congress for both chambers. Numbers of nodes are predominantly constant over time in both chambers. Number of edges increase from 1519 in congress 97 to 13247 in congress 117 in House of Representatives. In the Senate, networks with zero edges are because SDSM is unable to create network.

Table 4. Minimum, median, and maximum number of nodes for both chambers

Chamber	*Min. no nodes*	*Median*	*Max. no nodes*
House	378 (93rd congress)	423	497 (113th congress)
Senate	85 (113th congress)	117	127 (95th congress)

Table 5. Minimum, median, and maximum number of edges for both chambers

Chamber	*Min. no nodes*	*Median*	*Max. no nodes*
House	1519 (97th congress)	4259	13247 (117th congress)
Senate	0 (more congresses)	1	66 (117th congress)

5.2 Weak Polarization

Figure 4 shows the visualisation of weak polarization. In Figure 4, **A - D** show networks of both chambers with minimal and maximal modularity scores. For the House of Representatives, this is the 93rd (n = 378) and 114th (n = 444) congress respectively (see **A** and **C**). For the Senate, this is the 111th and 109th congress respectively (**B** and **D**). The blue nodes represent democrats, the red nodes the republicans, and the green nodes the independents. The edges shown in grey are the result of applying the unsigned SDSM, thus showing positive relations between legislators only. Nodes that are not connected to any other node did not (co-)sponsor more bills than at random together with another legislator. Figure **A** shows blue and red nodes being mingled, thus showing no clear community structures. Thus, there is minimal weak polarization within these networks. Figure **B** predominantly shows nodes without edges. There is only one edge, being between an independent and republican. Therefore, little can be said about the existence of weak polarization within this network. Figure **C** shows very clear distinction between red and blue nodes, showing two very clear community structures. This means that there is maximal weak polarization within these networks. Figure **D** predominantly shows nodes without edges. There are a few independents and some democrats being linked with each other. This could imply the existence of maximal weak polarization, but it could also be a coincidence. In Appendix A there is an overview for the networks of each congress.

In **E** and **F**, it can be observed that modularity index significantly monotonically increases over time following the Spearman's Rank Correlation Coefficient for the House of Representatives, but not for the Senate (House: $\rho = 0.56$, $p < 0.01$; Senate: $\rho = -0.29$, $p = 0.15$). The minimum value of modularity for the House of Representatives is $\rho = 0.0093$ in congress 93 and the maximum value is $\rho = 0.47$ in congress 114. Hence, the networks of congresses 93 and 114 for the House of Representatives are shown (**A** and **C**). The minimum value of modularity score for the Senate is $\rho = -0.5$ in congress 111 and the maximum value is $\rho = 0.46$ in congress 109. Hence, the networks of congresses 111 and 109 for the Senate are shown (**B** and **D**). Table 6 shows an overview of the modularity scores for both chambers with the p-values, and the minimum and maximum value of modularity scores.

Given that the modularity scores for all congresses in the House of Representatives are greater than zero, shows evidence for weak polarization in each congress. The modularity scores increase significantly over time, i.e., the extent of weak polarization has grown from minimal weak polarization in the 93rd congress, to maximal weak polarization in the 114th congress. This is supported by the visualisation of the network of the 93rd congress, showing no clear clusters,

and the network of the 114th congress, showing two very clear clusters. The one cluster exists mainly of republicans whereas the other cluster exists mainly of democrats. Since the modularity scores are calculated based on party affiliation, it shows that the partition matches the actual structure of the network.

Table 6. Modularity Score, p-value, minimum, and maximum value of modularity scores for both chambers

Chamber	***Modularity Score (ρ)***	***p-value***	***Min***	***Max***
House	0.56	< 0.01	0.0093	0.4745
Senate	-0.29	0.15	-0.5	0.46

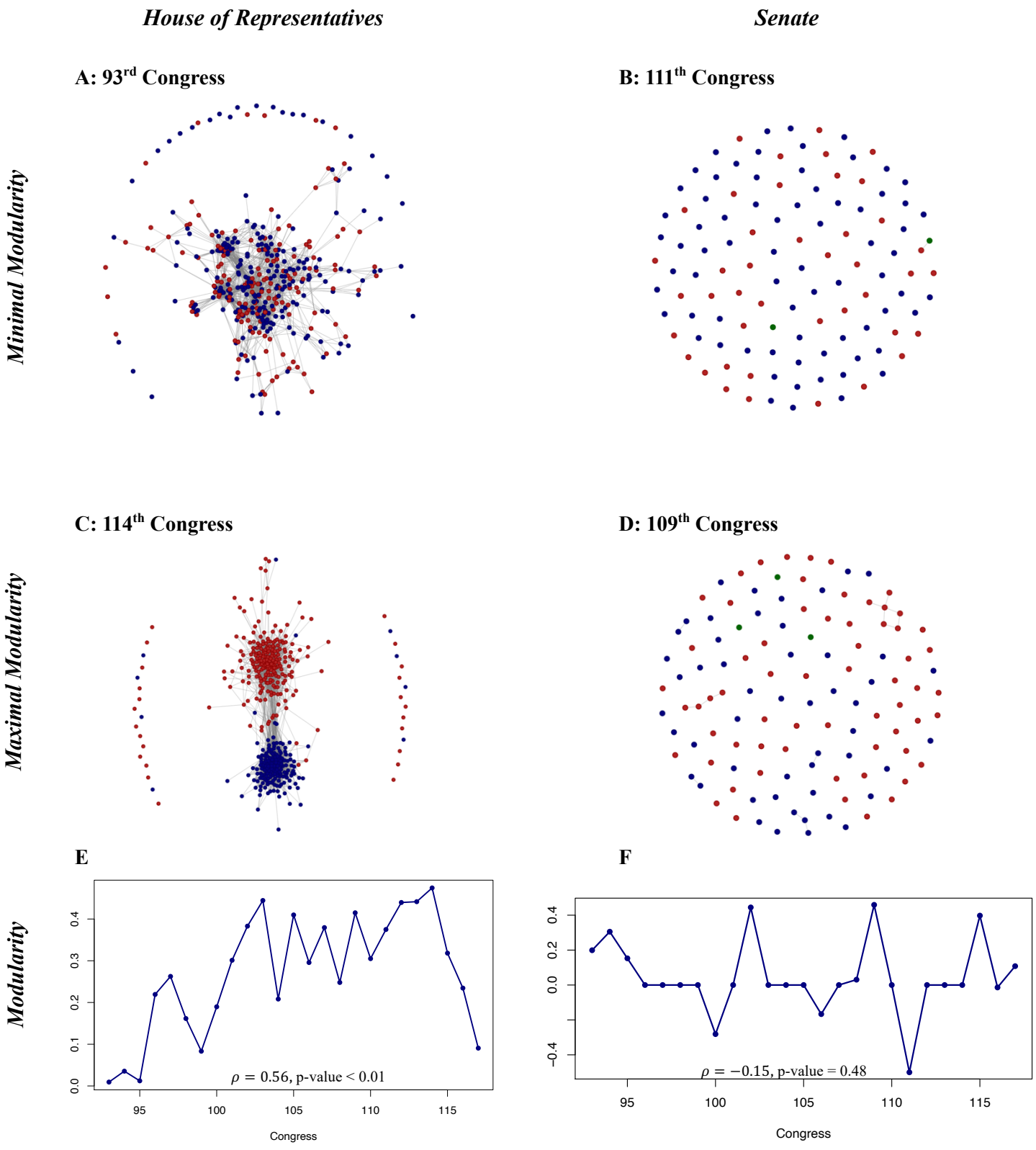


Figure 4. Visualisation of weak polarization. **A-D** show networks. Nodes are legislators with republicans in red, democrats in blue, and independents in green. Edges in grey are positive relations between legislators because of applying unsigned SDSM. **A** and **C** show the network of 93rd and 114th House of Representatives respectively. **B** and **D** show the network of the 111th and 109th Senate respectively. **A** shows one community structure, thus showing minimal weak polarization. **C** shows two clear communities, thus showing maximal weak polarization. **E** and **F** show modularity scores over time in House of Representatives and Senate respectively. It shows significant monotonically increasing trend in House of Representatives, but not in the Senate. Thus, weak polarization increases over time in the House of Representatives, but not in the Senate.

5.3 Strong Polarization

The number of nodes for the networks in the Senate are considerably low (see Figure 3B). As a result, there are no edges in a network for any congress when applying the signed SDSM. Hence, for the Senate it is not possible to show visualisation of strong polarization.

Figure 5 shows the visualisation of strong polarization. In Figure 5, **A** shows the Triangle Indexes over time in the House of Representatives. Even though it can be observed that the trend is negative, the result is insignificant (ρ = -0.19; p = 0.37). Values range between 0.75 and 1, indicating occurrence of more balanced triangles than would have been expected at random (0.75 > 0.5). The high values of the Triangle Indexes suggest that the proportion of unbalanced triangles (++- and ---) are almost negligible. Additionally, the scores suggest the existence of strong polarization in the House of Representatives over time, but nothing can be said about an increase or decrease in strong polarization over time by solely looking at the Triangle Indexes.

In Figure 5, **B** shows the decomposition of the Triangle Index, which are the two types of balanced triangles, i.e., +++ and +-- triangles, over time in the House of Representatives. The spearman correlation between proportion of +++ triangles and time show a significant monotonically decreasing trend in both the House of Representatives (ρ = -0.40; p < 0.05). Values range between 0.5 and 1, indicating occurrence of more +++ triangles than would have been expected at random (> 0.5). The spearman correlation between proportion of +-- triangles and time show a significant monotonically increasing trend in the House of Representatives (ρ = 0.46; p < 0.05). Values range between 0 and 0.5, indicating occurrence of +-- triangles is less than would have been expected at random (< 0.5). The decrease of +++ triangles and increase of +-- triangles show a shift from positive balanced triangles (+++) towards hostile balanced triangles (+--). This supports the idea that strong polarization increases over time.

In Figure 5, **C** shows the proportion of the +++ triangles from **B** that are within party and shows the proportion of the +-- triangles that are between party. Figure 6 shows an overview of the different types of +++ triangles that are within party and +-- that are between party: +++ triangles within party always exist of either three democrats linked to each other or three legislators. The +-- triangles between parties always exist of either two democrats and one republican, in which the positive relation is between the two democrats, or of two republicans and one democrat, in which the positive relation is between the two republicans. It can be observed that there is a significant monotonically increasing trend for the +++ triangles within parties (ρ = 0.12; p < 0.05). There is also a significant monotonically increasing trend for the +-- triangles between parties (ρ = 0.64; p < 0.05). Table 7 shows an overview of the results mentioned above.

Additionally, it can be observed that approximately 50 percent of the +++ triangles exist within party in each congress, sometimes increasing to 80 percent in both chambers. Additionally, it shows that before the 105th congress, approximately 20 percent of the +-- triangles exist between parties and after the 105th congress, approximately 40 percent of the +-- triangles exist between parties in both chambers. There is an approximate increase from 66 percent from the minimum proportion of +++ triangles to the maximum proportion of +++ triangles. Also, there is an approximate increase of 400% in proportion between the minimum and maximum +-- triangle between parties. The stronger increase of +-- triangles between party

than +++ triangles within party further supports the idea that strong polarization increases over time.

Thus, even though the Triangle Indexes do not support the existence of an increase of strong polarization over time, breaking down the Triangle Indexes in the two distinct balanced triangles and decomposing them even further into +++ triangles within party and +-- between party do support an increase in strong polarization.

Table 7. Spearman Correlations and p-values of Triangle Index, +++ and +-- triangles, +++ triangles within party, and +-- triangles between party

	Spearman Correlation (ρ)	***p-value***
Triangle Index	-0.19	0.37
+++ triangles	-0.40	< 0.05
+-- triangles	0.46	< 0.05
+++ w/in party	0.12	< 0.05
+-- btw party	0.64	< 0.05

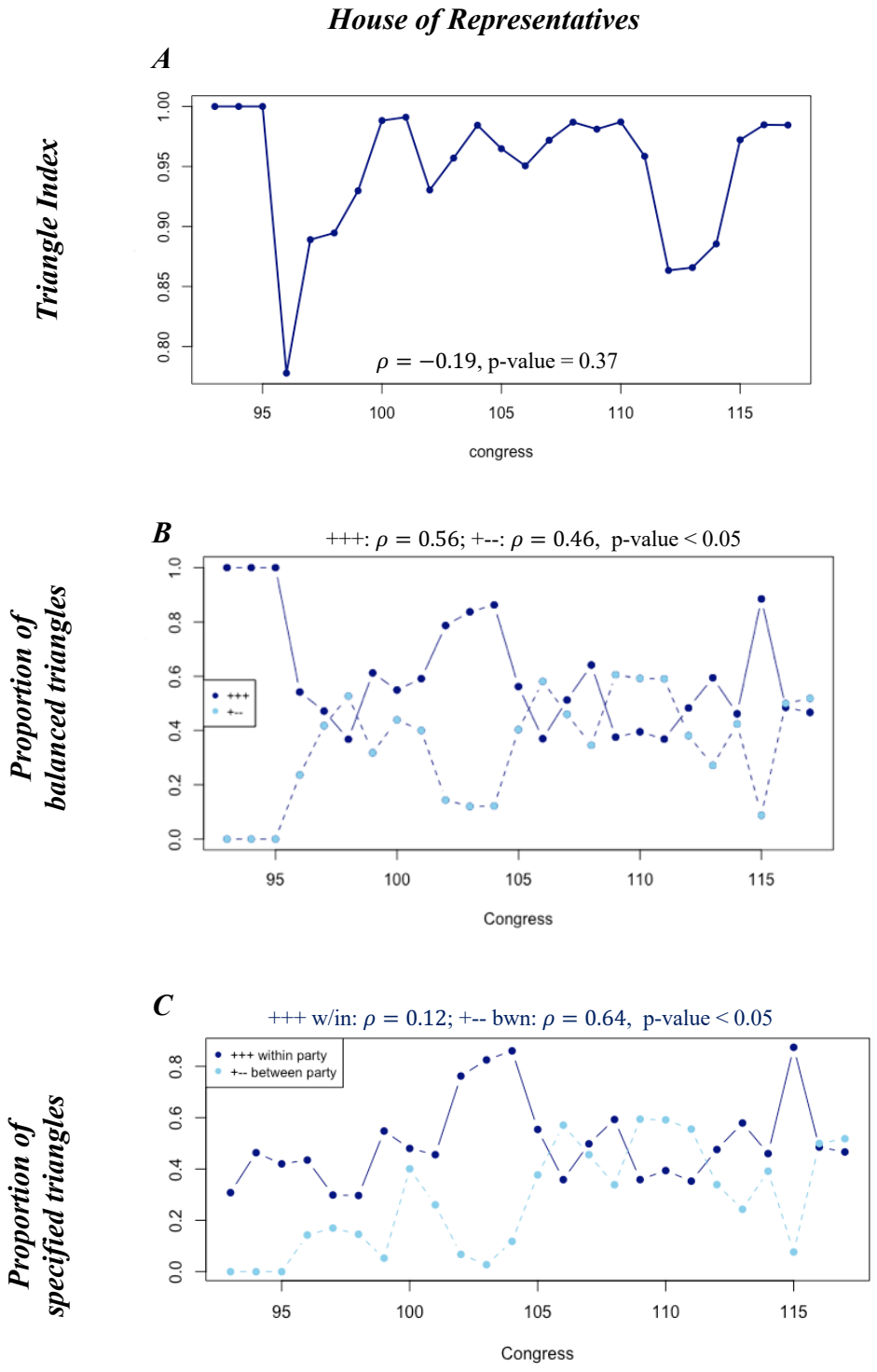


Figure 5. Visualisation of strong polarization. **A** shows the Triangle Index for the House of Representatives. There is no significant trend of the Triangles Indexes over time. **B** shows the decomposition of balanced triangles, consisting of +++ and +-- triangles. The +++ and +-- triangles show a significant decreasing and increasing trend respectively. The +++ triangles are more common than the +-- triangles. **C** shows the proportion of +++ triangles within the party (D or R) and +-- between the parties. It shows that approximately 50% of the +++ triangles exist within party in each congress, sometimes increasing to 80% in both chambers. Additionally, it shows that before the 105th congress, approximately 20% of the +-- triangles exist between parties and after the 105th congress, approximately 40% of the +-- triangles exist between parties in both chambers. The results show evidence for existence and increase of strong polarization over time in House of Representatives.

Possible +++ triangles within party ***Possible +-- triangles between party***

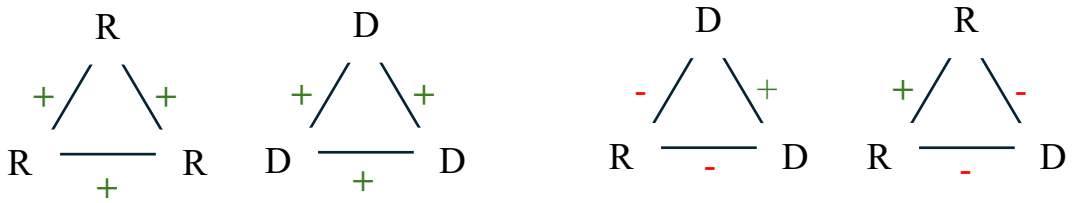


Figure 6. Visualisation of possible +++ triangles within party and +-- between party. The +++ triangles are always between three republican or three democratic legislators. The +-- triangles always consist of either one republican and two democrats or two republicans and one democrat. The positive relation is always between the two legislators of the same party and the two negative relations between the legislators of opposing party.

6 Discussion

6.1 Main Findings

This master thesis assessed how political polarization manifests in the United States Congress from the 93rd to the 117th sessions, and to what extent there is evidence of ideological and affective polarization in the interactions between democratic and republican legislators. Supported by literature, the underlying hypothesis was that weak polarization, as a possible measure of ideological polarization, would not only be present within each congress in both chambers, but also increase over time. Additionally, strong polarization, as a possible measure of affective polarization, would not only be present within each congress in both chambers, but also increase over time. The main finding of this master thesis is that weak polarization increases significantly over time in the House of Representatives, but not in the Senate. Additionally, there is no clear consensus regarding whether strong polarization increases over time in the House of Representatives and remains unclear for the Senate.

6.2 Interpretation of the Findings

6.2.1 Weak Polarization

Findings regarding the existence of weak polarization in the House of Representatives are considerably consistent, as shown in the results. There is a significant monotonically increase of modularity scores over time in the House of Representatives, which is in line with the results of Neal (9). The modularity scores reach its maximum at the 114th congress and decreases after, which implies a decrease in cluster formation from the 115th congress onwards. This is not in line with the visualisation of the 115th, 116th, and 117th congress, that each show two very clear distinct clusters (see Appendix A). However, there can be multiple reasons for this decrease in modularity. Firstly, modularity scores are also influenced by clustering within clusters. If there are strong clusters within the big clusters, the modularity score can be lower since it possibly emphasizes the network structure of the smaller clusters. Even though this could be the case for the 115th, 116th, and 117th congress, the networks do not show clear clusters within clusters, so this is not too likely. In turn, modularity scores are also influenced by the density of the edges within each cluster, i.e., how many edges there are relative to the number of possible edges. If the density within each cluster is very high, the modularity score can decrease. Indeed, congress 115, 116, and 117 show very dense clusters, being a possible reason for the decrease in modularity score. The evidence for weak polarization and an increase in weak polarization over time in the House of Representatives are an indication for the existence and increase in ideological polarization over time in the House of Representatives.

Findings regarding the existence of weak polarization in the Senate are not significant. Therefore, no conclusions can be drawn regarding the existence or increase of ideological polarization over time in the Senate. The results can be insignificant for multiple reasons. The number of legislators in the Senate are relatively small, as shown above. Therefore, the number of possible edges is relatively low, which can influence the flexibility of the model being applied (SDSM in this case). Secondly, the degree sequences of the SDSM, i.e., the degree of each node in the network, are very specific for a small network. Therefore, it becomes difficult for the model to create a distribution. Hence, it becomes even more difficult to create edges in the

network, as these are based on that distribution. The modularity scores can be influenced by these small networks, since the scores are sensible to noise and overfitting.

6.2.2 Strong Polarization

Findings regarding the existence of strong polarization in the House of Representatives are inconsistent, as shown in the results. There is no significant spearman correlation of the Triangle Index over time. On the contrary, the decomposition of the balanced triangles (+++, ++-), i.e., the decomposition of the Triangle Index, show a significant decreasing trend for the +++ triangles and a significant increasing trend for the ++- triangles, as shown above. These results are not in line with Neal, who showed significant increase of the Triangle Index over time and significant increase of both the +++ and ++- triangles over time (9). However, this could be the result of me using my own dataset of bills, inherently creating different networks despite both applying the same signed SDSM. The +++ triangles within party and the +-- triangles between party show a significant increasing trend, and are in line with the results of Neal (9). These significant results could indicate that affective polarization is increasing within the House of Representatives. However, these conclusions should be drawn with much caution, since the Triangle Index scores are insignificant.

Findings regarding the existence of strong polarization in the Senate are not existing. Therefore, no conclusions can be drawn regarding the existing of affective polarization in the Senate. The absence of results for the Senate are the direct consequence of the signed SDSM not being able to create a network for such small number of nodes, as mentioned above.

6.3 Strengths and Limitations

A strength of this thesis is that the same method is used as applied by Neal, but with using an own set of data instead of the built-in dataset in the *Backbone* library of Neal (9). This means that it is also possible to apply the methods by Neal within parliaments of other countries than the United States. Another strength of this thesis is that the datasets of the House of Representatives are considerate enough to create significant results.

On the other hand, a limitation of this thesis is that the datasets of the Senate are not considerate enough to create significant or even any results. Another limitation is that results stem from congress 93 and onwards. It would have been more insightful if data was included from congress 1 and onwards. However, congress 93 was chosen deliberately, because of the lack of information of congresses before that time. Lastly, the SDSM is a model based on relatively simple statistical assumptions about the degree distribution. Therefore, it can potentially have missed realistic characteristics of the actual networks.

6.4 Implications and Recommendations

This thesis measured ideological and affective polarization by considering weak and strong polarization as indicators of ideological and affective polarization. However, as these are only indicators and these relations are not one-to-one, more research could be done in researching other indicators of ideological and affective polarization. Then, those results could be combined with those of the weak and strong polarization. This leads to more adequate quantification of ideological and affective polarization.

7 Conclusion

The consequences of political polarization can be both pernicious as well as beneficial. Different forms of political polarization, such as ideological and affective polarization explain how political polarization can move towards a pernicious or beneficial state. Much research has already been conducted on quantifying ideological polarization. Calculating DW-NOMINATE scores is a widely accepted method for quantifying ideological polarization by placing individuals on a scale from conservative to liberal. However, the inability of quantifying affective polarization through DW-NOMINATE scores gave rise to other methods of quantifying political polarization. Considering political polarization as a network science problem allowed for quantifying both ideological as well as affective polarization through determining weak and strong polarization, which are considered indicators of ideological and affective polarization respectively. Weak polarization was measured by estimating modularity scores. Strong polarization was measured by determining the degree of balance of the network. Modularity scores decreased significantly over time in the House of Representatives, indicating a decrease in clustering over time, which was in accordance with the visualisation of the networks. There was no significant trend of the Triangle Index over time. However, the decrease of +++ triangles and increase of +-- triangles were significant. Additionally, the increase in both +++ triangles within party and +-- triangles between party were also significant. Therefore, the results remain inconsistent, and little can be said about the increase of affective polarization over time. More research is needed to further develop a working method to examine occurrence of ideological and affective polarization.

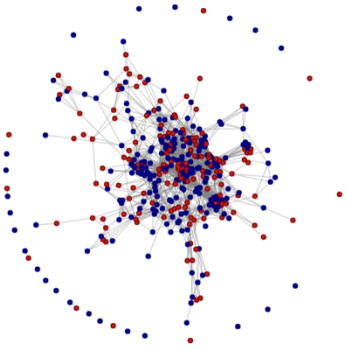
Bibliography

1. Schedler A. Rethinking Political Polarization. Political Science Quarterly. 2023 Aug 31;138(3):335–59.
2. Fiorina M, Abrams S. Political Polarizaion in the American Public. Annual Review of Political Science. 2008 Jun 10;11.
3. Heltzel G, Laurin K. Polarization in America: two possible futures. Curr Opin Behav Sci. 2020 Aug;34:179–84.
4. Tucker JA, Persily N, editors. SSRC Anxieties of Democracy. In: Social Media and Democracy [Internet]. Cambridge: Cambridge University Press; 2020 [cited 2024 Jun 16]. p. ii–ii. (SSRC Anxieties of Democracy). Available from: https://www.cambridge.org/core/books/social-media-and-democracy/ssrc-anxieties-of-democracy/FBB0E319323C4BEEC3590D504642101C
5. Bell P. Public Trust in Government: 1958-2023 [Internet]. Pew Research Center. 2023 [cited 2024 Jun 16]. Available from: https://www.pewresearch.org/politics/2023/09/19/public-trust-in-government-1958-2023/
6. McCoy J, Somer M. Toward a Theory of Pernicious Polarization and How It Harms Democracies: Comparative Evidence and Possible Remedies. The ANNALS of the American Academy of Political and Social Science. 2019 Jan 1;681(1):234–71.
7. Weber TJ, Hydock C, Ding W, Gardner M, Jacob P, Mandel N, et al. Political Polarization: Challenges, Opportunities, and Hope for Consumer Welfare, Marketers, and Public Policy. Journal of Public Policy & Marketing. 2021 Apr 1;40(2):184–205.
8. Axelrod R, Daymude JJ, Forrest S. Preventing extreme polarization of political attitudes. Proc Natl Acad Sci U S A. 2021 Dec 14;118(50):e2102139118.
9. Neal ZP. A sign of the times? Weak and strong polarization in the U.S. Congress, 1973–2016. Social Networks. 2020 Jan 1;60:103–12.
10. Lewis, Jeffrey B., Keith Poole, Howard Rosenthal, Adam Boche, Aaron Rudkin, and Luke Sonnet. Voteview: Congressional Roll-Call Votes Database [Internet]. 2024 [cited 2024 Jun 16]. Available from: https://voteview.com/data
11. Download Search Results [Internet]. [cited 2024 Jun 16]. Available from: https://www.congress.gov/help/using-data-offsite/download-search-results
12. Leeper TJ. The Informational Basis for Mass Polarization. Public Opinion Quarterly. 2014 Jan 1;78(1):27–46.
13. Enders AM. Issues versus Affect: How Do Elite and Mass Polarization Compare? The Journal of Politics. 2021 Oct;83(4):1872–7.
14. Banda KK, Cluverius J. Elite polarization, party extremity, and affective polarization. Electoral Studies. 2018 Dec 1;56:90–101.
15. Zingher JN. Mass Opinion in Context: How Elite Polarization Has Shaped Political Behavior. In: Zingher JN, editor. Political Choice in a Polarized America: How Elite Polarization Shapes Mass Behavior [Internet]. Oxford University Press; 2022 [cited 2024 Jun 16]. p. 0. Available from: https://doi.org/10.1093/oso/9780197630693.003.0002
16. Poole KT, Rosenthal H. The Polarization of American Politics. The Journal of Politics. 1984 Nov;46(4):1061–79.
17. DeSilver D. The polarization in today's Congress has roots that go back decades [Internet]. Pew Research Center. 2022 [cited 2024 Jun 13]. Available from: https://www.pewresearch.org/short-reads/2022/03/10/the-polarization-in-todays-congress-has-roots-that-go-back-decades/
18. Tajfel H, Turner JC. The Social Identity Theory of Intergroup Behavior. In: Political Psychology. Psychology Press; 2004.
19. Heider F. The psychology of interpersonal relations. Hoboken, NJ, US: John Wiley & Sons Inc; 1958. ix, 326 p. (The psychology of interpersonal relations).

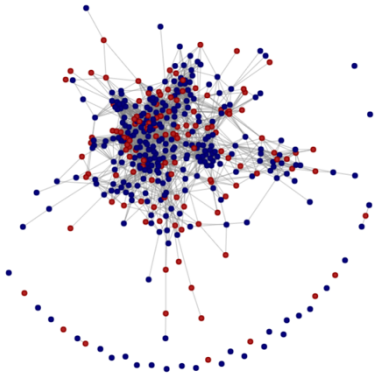
20. Voteview | Data [Internet]. [cited 2024 Jun 28]. Available from: https://voteview.com/data
21. ProPublica Data Store [Internet]. [cited 2024 Jun 28]. ProPublica Data Store. Available from: http://www.propublica.org/datastore/datasets
22. Woon J. Bill Sponsorship in Congress: The Moderating Effect of Agenda Positions on Legislative Proposals. The Journal of Politics. 2008 Jan;70(1):201–16.
23. Schiller WJ. Senators as Political Entrepreneurs: Using Bill Sponsorship to Shape Legislative Agendas. American Journal of Political Science. 1995;39(1):186–203.
24. Wilson RK, Young CD. Cosponsorship in the U. S. Congress. Legislative Studies Quarterly. 1997;22(1):25–43.
25. Newman MEJ. Modularity and community structure in networks. Proceedings of the National Academy of Sciences. 2006 Jun 6;103(23):8577–82.
26. Clauset, A.; Newman, M. E. J. & Moore, C. Finding community structure in very large networks. (Physical Review E 2004, 70, 066111).
27. Spearman C. The Proof and Measurement of Association between Two Things. The American Journal of Psychology. 1904;15(1):72–101.
28. Structural Balance [Internet]. [cited 2024 Jun 16]. Available from: https://cran.r-project.org/web/packages/signnet/vignettes/structural_balance.html
29. Schoch D. signnet: An R package for analyzing signed networks. Journal of Open Source Software. 2023 Jan 27;8(81):4987.

Appendix A

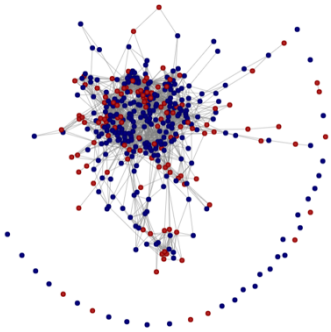
Networks for congresses 93 until 117



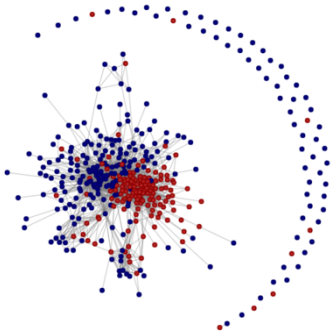
93



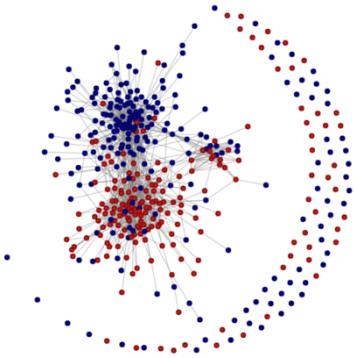
94



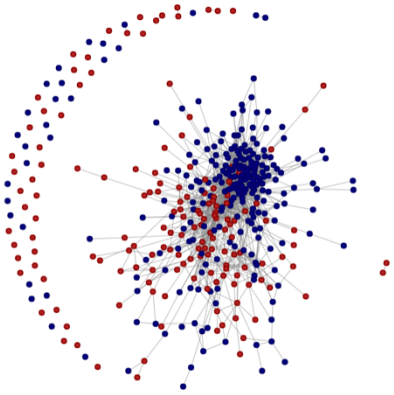
95



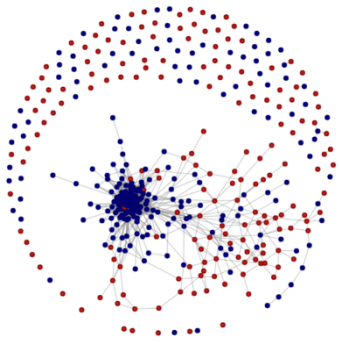
96



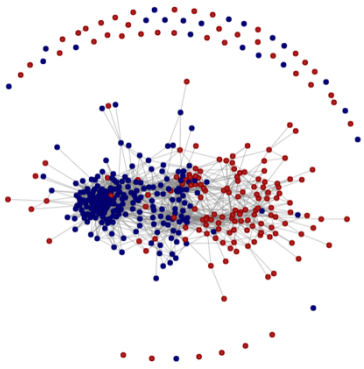
97



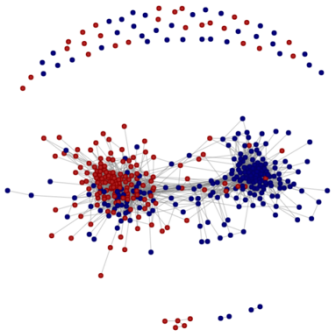
98



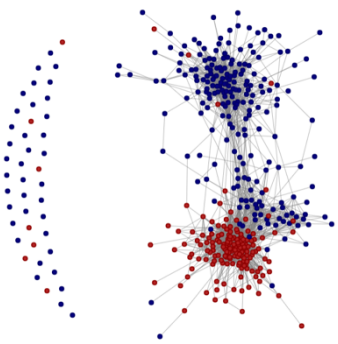
99



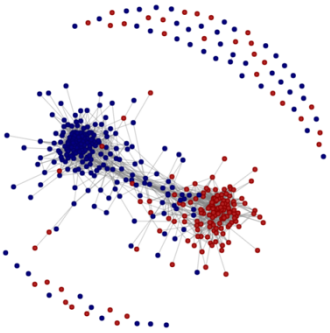
100



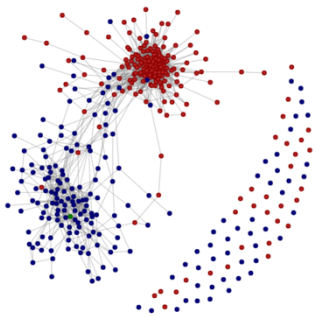
101



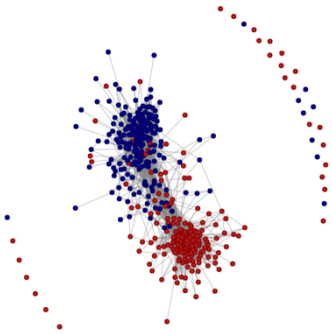
102



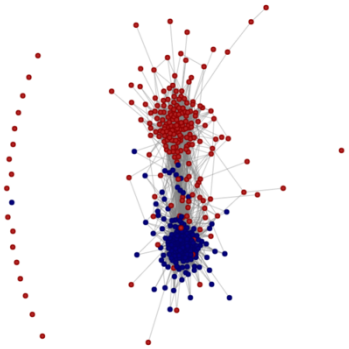
103



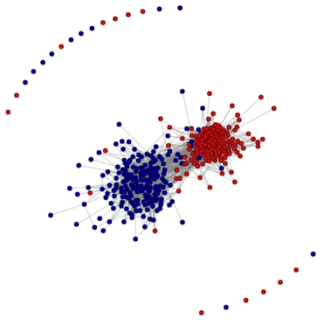
104



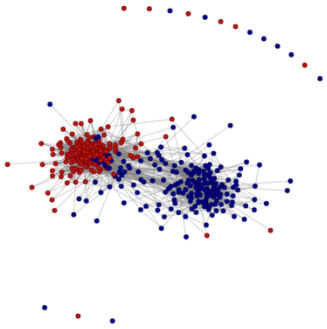
105



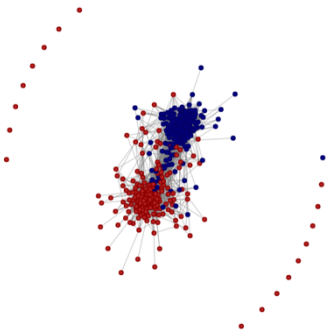
106



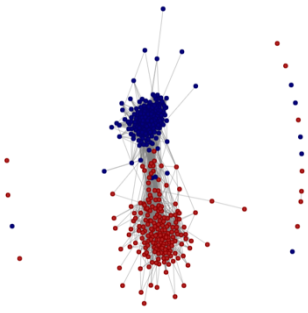
107



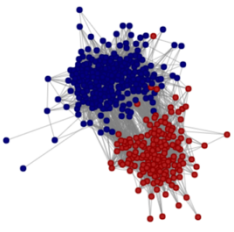
108



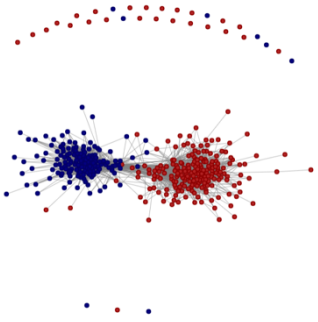
109



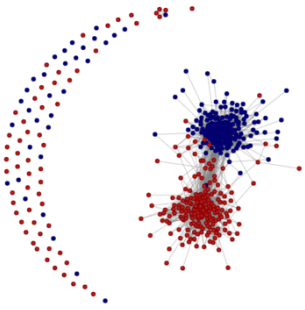
110



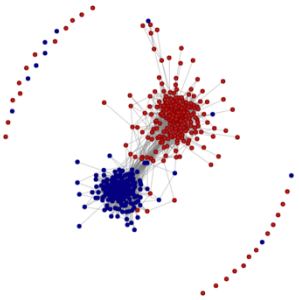
111



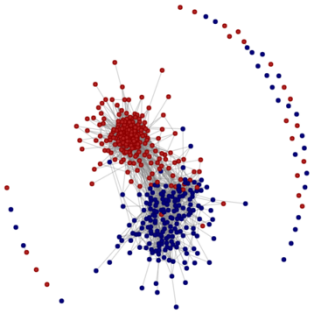
112



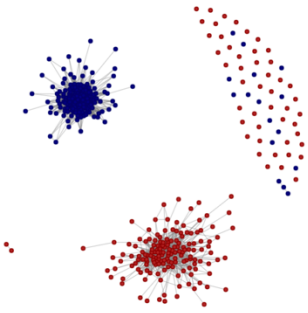
113



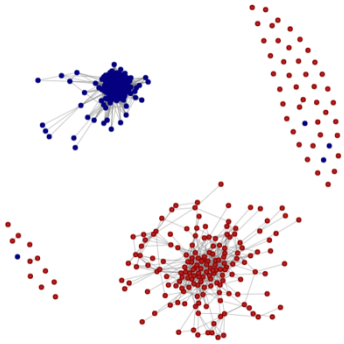
114



115



116



117