Polarization in U.S. political community

Qualitative and Quantitative Methods (BDMAO1026U)

Ralf Joosep Kask

Pages: 10

Characters: 32051

Hand-in Date: 2024-06-06

COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

Contents

1 Introduction

As we are living at the era of information, big data and larger processing storage and speed enable information creation and sharing to flourish. However, as with every good tool, this can be used badly and create problems. In particular, disinformation and information bubbles add fuel to the polarization, which is deemed also as one of 21st century's great challenges (Potgieter & Cooper, 2022). While polarization of society is high due to social media, it can be argued that media and politicians can be root cause of burning flames. Therefore, to give some backing to the argument and whether polarization and to what extent is a problem, it is important to explore the area.

Polarization can occur in social networks with high interconnectedness of subcommunities yet weak inter-group connectivity. As Interian et al. put it, some polarization can be good, especially in political debate instances of democratic societies. The further, extreme polarization can however lead to undesired results with possibility of overthrowing peaceful democratic way of life. In their paper dealing with network polarization measures, they lay a theoretical foundation of this paper of trying to grasp change of polarization in U.S. congress debate. (Interian et al., 2023)

Interian et al. discover that behind any polarization model is concept of groups, i.e. the fragmentation into groups based on sharing common characteristics. In particular, they identified measures such as homophily, modularity, random walk controversy, content qualification methods and balance-based measures. According to Interian et al., homophily determines tendency to be associated with more similar ones based on characteristics such as degree or other attributes, modularity evaluates "the number of intra-community against inter-community edges for a given set of node groups in a network". The other three are more complex and require more theory. (Interian et al., 2023)

This paper aims to apply homophily and modularity as measures to estimate polarization change occuring at U.S. Congress. The paper does so by building a graph of congress members through connecting members by topics occurring in texts. Therefore, we arrive at a research question this papers tries to answer: How has the polarization of the U.S. Congress political community changed from 1970 to 2020? As this paper is exemplary study, not full-fledged research, we only carry analysis out on data segment from July 20, 2005 and therefore are not able to give any real answers about the change of polarization, however it can be arrived if process is carried out as outlined in the following sections using entire data.

The paper is divided further into 4 sections. The next section highlights the methodology. It is followed by carrying out the methodology in data analysis section. Discussion entails reflective view on used approach and other possibilities and final section of conclusion wraps up the paper.

2 Methodology

The chosen methodological approach takes quantitative angle. In particular, we utilize combination of machine learning and network analysis techniques in order to understand the underlying community structure of political debate held at U.S. congress. The main methodology framework follows already demonstrated on New Zealand's parliamentary speeches (Curran et al., 2017).

Curran et al. in their paper examine how political activity and topic popularity changes over time by implementing "a Latent Dirichlet Allocation model to discover the thematic structure" and then building network consisting of Members of the Parliament from "two-mode networks linking Members of the Parliament to the topics they discuss" (Curran et al., 2017). In nearly identical fashion, we carry out the preparation of the network in a same way. The analysis consists by utilizing community structure detection based on edge betweenness and calculating assortativities.

2.1 Top Modelling

In order to understand free-form string data, text data, there needs to be a step convert it to more discernible. One way to categorize data is to perform topic modelling. While topic modelling can be done through qualitative coding, Latent Dirichlet Allocation (LDA) has proved to be valid and most used topic modeling technique for vast amounts of texts. On top of previously mentioned benefit of automated content coding, LDA allows for polysemy (one word with two or more distinct meanings) because topics are not mutually exclusive, i.e. individual words appear across topics with differing probabilities, and topics themselves may overlap or cluster. (Storopoli, 2019)

While not going to deep how LDA works, on surface level, LDA assumes that documents with similar topics use similar words and uses bag-of-words approach, i.e. uses individual units and does not consider relationship of units. LDA generates topic probability distribution for each document (Maklin, 2022). LDA requires crucial hyperparameter of K, which stands for number of topics LDA has to learn. Generally, it is wise to create a model for each value of K and find the best fit assessing the generated either quantitatively (using some algorithm) or qualitatively (based on researches judgements) (Storopoli, 2019). In this paper, for exemplary purposes, we set $K = 30$ as it seemed to be valid, used by Curran et al. and many papers settling in range of 5 to 50 and with median at 20 (Fišer & Pahor de Maiti, 2021). The LDA results in a topic-word matrix and in topic-document matrix, where former is a matrix where each topic has weights of words associated with the topic and latter is a matrix where each document weighs topics for the document (Storopoli, 2019).

2.2 Social Networks

After texts have been extracted, the process follows on by creating social networks. Social networks can be formed by composing or aggregating relations between different units. In social network, one unit can be called social actor or mathematically node. A relation which links two different actors is called mathematically an edge. In network analysis, it is not individuals who are interest of the study, it is relationship and structure as a whole. (Martino & Spoto, 2006)

More mathematically, a network or a graph G is an pair (V, E), where V is a set of vertices (nodes) and E is a set of edges (Gordon College, n.d.). A bipartite graph, is a graph for which nodes can be decomposed into two separate sets so that no two graph vertices in the same set share an edge (Weisstein, n.d.). Or in other words, bipartite graph G is an triplet (U, V, E), where U, V are disjoint sets and E set of edges not containing edges connecting nodes from separate sets (Curran et al., 2017).

In our case, first by forming bipartite graph G, we set U to be our learnt topics, V to be members of congress and edges E to be all linkages between people and topics. These linkages we define to be if member has a text classified with particular topic, then there is an edge between that member and that topic. Of course, as Curran et al. did by introducing lower bound for edge creation of 6.7% of speeches in a year, such criteria of edge formation likely needed when data taken under consideration spans over multiple years as Curran et al. did, in order to carry out proper full-fledged research (in our case we only have one single debate day).

After bipartite (two-mode) structure, we transform the structure to one-mode projection as we are more interested in members of congress and their relations. Since, one-mode projection is often less informative than two-mode network (as we drop information), a weighting of nodes can retain the original information better (Zhou et al., 2007). Indeed, projection keeping weighting can be achieved through matrix multiplication - for overlap count (Murphy & Knapp, 2018). It is easy to convert bipartite matrix B to one mode matrices A through $A = B \times B^T$. In this way, we use simple weighting method, where now graph has only members of congress as nodes and edges with weights, where edge weight is equal to number of shared topics between the neighboring members. Weighting is important in our case as community detection is challenging in the absence of weighting due to possibility of popular topics (Curran et al., 2017). Figure 1 shows the process of transforming bipartite to one-mode projection.

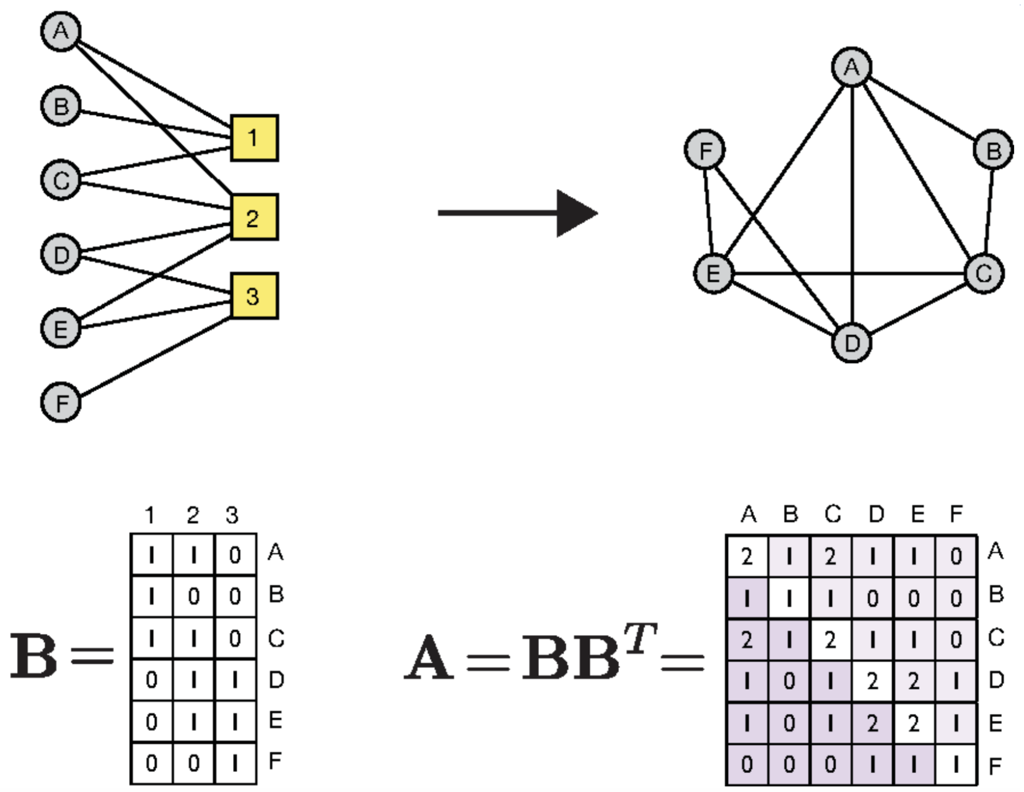

Figure 1: Transforming bipartite to one-mode projection, from O'Malley and Onnela, 2014

3 Data analysis

Data is extracted from U.S. Congressional Records (Congress.gov, 2005). The chosen data is from a single day of debate at the United States Congress on July 20, 2005. In particular, data set consists of dataframes of legislators and records. Of course, in proper research paper, data collection is carried out systematically and this particular data instance is only used to serve as an example how data analysis would be done in actual case.

Data analysis is carried out using tool R, which is well-known programming language for statistical computing (R Core Team, 2024). The topic modelling is carried by using R packets `rJava` and `mallet` and graph creation is created by packet `igraph`. The used code can be found in full length in appendix A.

Firstly, in order to ensure relevant cleaned data, we dismiss and drop bill labels and therefore keep only unique texts to be taken under analysis as data consists of duplicate texts under different bill names. Secondly, before moving on to one of the most important step of analysis, Mallet and Latent Dirichlet Allocation, we filter out texts that are smaller than 20. This is due to small texts do not carry any real political content and are procedural statements rather. In full-fledged research, it is likely to raise the filter lower bound in order to ensure that all texts taken under consideration contain actual themes, not like the following, which we e.g. filtered out: "Mr. Chairman, I offer an amendment." - longer the text, more likely it contains relevant information and thus political content. Also, as texts contain new line characters `\n`, we also do simple pre-processing by removing those characters.

Moving on to R Mallet and Latent Dirichlet Allocation, we perform instance import with English stopwords (removing words which universally does not carry information). Mallet model import has several hyperparameters, which need to be tweaked in order to optimize the model performance. In our case, we opt for defaults and number of topics $K = 30$ as already written in previous section. Also, in order to ensure reproducibility, crucial part is to include `setRandomSeed(42L)`.

After the model has done its job, as already mentioned, model gives probabilistic distributions for every document and we for simplicity we narrow the topics of text to single concrete topic by assigning each document only the most topic with highest share.

Then, considering bipartite incidence matrix and converting it to one-mode incidence matrix, we are ready to plot the graphs and perform analysis. Figure 2 displays the resulting graph with coloring nodes based on party belonging.

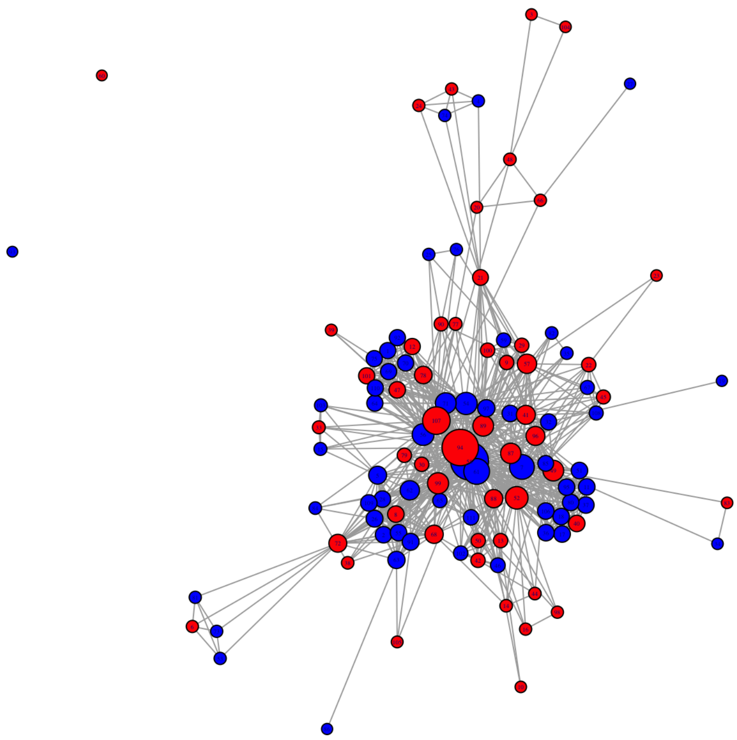

Figure 2: Graph

In order to observe the change, such graphs can be done iteratively and then compared as time progresses. Time intervals must relatively long in order to observe fundamental change, e.g. yearly, two-year terms or six-year terms, where last two correspond to length of terms of House of Representatives or Senators respectively (Walberg, n.d.).

Metrics we consider, according to the theory, are modularity and homophily. After applying community structure detection based on edge betweenness, modularity can be calculated and observed over the course of time by plotting a time series chart. When it comes to homophily, this we can calculate using assortativity by class or degree assortativity. These can be also observed over the course of time and then plotting a time series chart for truthful, beautiful and informative overview of change. Community structure detection recognizes following clusters as shown in figure 3.

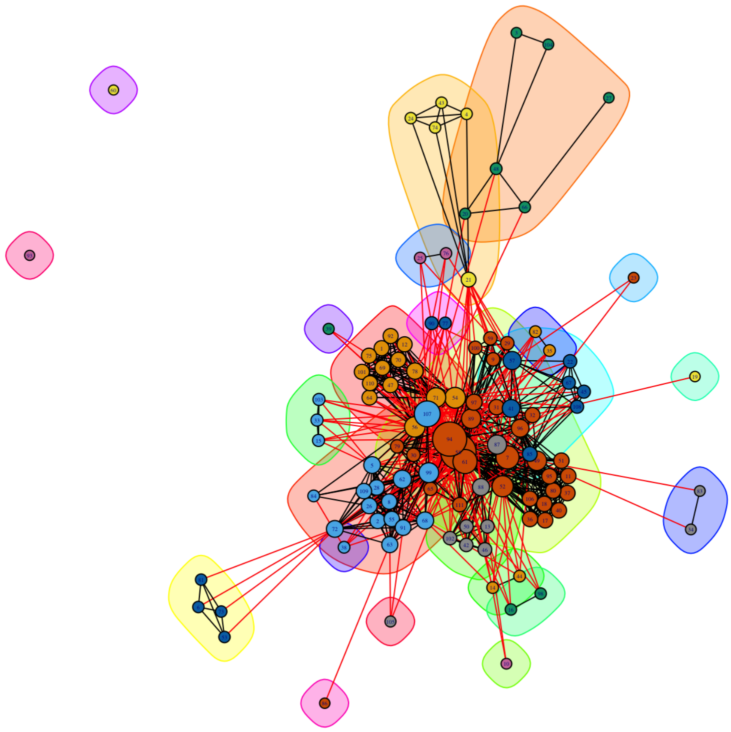

Figure 3: Graph with community structure detection based on edge betweenness

Exemplary results are shown in table 1. Taking a quick look at the results, we observe that modularity of 0.35 tells us some modularity (-1 would be almost complete network and 0 modularity of random graph), homophily for both by party and degree falls near 0, yet corresponding with results from Curran et al. However, it is important to remember that these stand-alone numbers do not tell us nothing about the change which was set out in the research question (again, in a proper study, one would consider timeframe of our research question and then can compare and note the change).

time	modularity	assortativity_party	assortativity_degree
2005-07-20	0.3462061	0.02949582	-0.1373556
⋮	⋮	⋮	⋮

Table 1: Table of results

4 Discussion

First of all, choice of methodology is highly dependent on research question paper tries to uncover. We try to understand how has polarization changed at U.S. Congress from 1970 to 2020. The time frame stated in the question is so vast therefore it is likely it requires quantitative approach. Also, when it comes to quantitative approach to science, objectivist and positivist view is associated with it and thus it seeks towards universal claims and gives universals same ontological status as particulars, i.e. possibilities to be included in scientific claims, which aligns with this paper.

Since concept of polarization is universal and not particular, like gravity and love, there needs to be some type of operationalization and this can be defined differently. We used paper from Interian et al. to be theoretical foundation of this work. Based on used theory and methodology to understand polarization in U.S. congress, the approach of this paper seeks to identify interrelated groups and measure the structure.

Using different theoretical framework, one might try to explore the area through use of surveys and thus not fall into textual data as Lelkes point out. Thus, taking the angle of ideological alignment or perceived polarization of opposite partisans, one may instead use surveys as well to seek clarity. (Lelkes, 2016)

Coming back to textual data and thus precisely to texts from congress, Peterson and Spirling used supervised machine learning and its performance to estimate polarization in parliamentary texts (Peterson & Spirling, 2018). In that sense, it measures relative change of polarization and does not try to frame polarization as a state. Also, apart from supervised and unsupervised (clustering) techniques, Catarina Pereira and Rosa also notes of building statistical/parametric models (analogy with content analysis) which are simpler than machine learning, e.g. "political polarization measurements are derived from the pro-portion of supporting terms" (Catarina Pereira & Rosa, 2024).

Other quantitative approaches such as experiment and longitudinal study are hard to be employed especially considering subject of study and time frame respectively deflating possibilities of those.

Our approach of LDA + network analysis (with 2 polarization measures) has strength of relative simplicity of actually measuring polarization based on the theoretical framework. One huge weakness of used methodology is edge formation as it uses premise of polarization groups formed by same topic bounds, on top of that, we made in the process assumption that each text is about only one topic and thus this gives weight to some edges and removes some edges at all. Other minus is about the two process of two steps which both blur term of polarization and increase error. As improvements one can find better hyperparameter tuning, better measures and limitation of the work is that it seems unnecessarily complex, yet is quite informative. Indeed, supervised machine learning and non-machine learning statistical and parametric models ensure reproducibility and are easier than unsupervised topic modelling and network analysis applied. They however might be too plain and measures of polarization not that concrete. Comparing those with completely different approach, survey data (from e.g. American National Election Studies) can be applied for perceived polarization or ideological alignment, this differs from textual data and thus is more inflexible and data lacks depth. Also, preparing and analysing surveys might be more time-consuming, than taking archival data (such as our U.S. congressial texts from Congress.gov). Table 2 categorizes the discussed.

methodological approaches	strengths	weaknesses	improvements	limitations
unsupervised LDA + network analysis (this paper) (Curran et al., 2017), (Interian et al., 2023)	simplicity, clear foundation of pre-existing frameworks	multiple steps, each can abstractify term of polarization	possibility of better LDA training, possibility of more in breadth and depth measurements	unnecessarily complex
survey data like American National Election Studies or Pew Research Center (Lelkes, 2016)	different theoretical framework, flexibility in terms of perceived polarization and ideological alignment scales	time-consuming, inflexibility		limitations in terms of survey setup (depth of understanding)
supervised machine learning (Peterson & Spirling, 2018)	reproducibility, simplicity	measure of polarization		fixed categories
statistical and parametric models, content analysis (Catarina Pereira & Rosa, 2024)	reproducibility, easier than machine learning	too plain, measurement of polarization most abstract, class attribution is manual	can be implemented automatic class attribution	

Table 2: Table of methodological angles

5 Conclusion

The paper took quantitive approach and tried to measure polarization in social network of congress after Latent Dirichlet Allocation algorithm of topic modelling. The chosen measures were modularity and homophily, more precisely modularity found by cluster edge betweenness algorithm, degree assortativity and assortativity by party based on theoretical framework of Interian et al. The chosen methodology was carried out in R, using R mallet for topic modelling and igraph for network creation and measurements.

While this paper did not find definitive answer to the research question, it outlined the methodology and showed this process on exemplary data from texts of U.S. congress held on July 20, 2005. Results corresponded with methodological framework of Curran et al.

The paper also gave ground for further research and improvements as it found other methodological approach and compared them to the used one. For example, while this paper used unnecessarily complex process with need of good training, statistical and parametric models with context analysis can be easier to understand, create and recreate. Supervised learning also has upper hand on simplicity, yet has unclear measure of polarization and limitation of fixed categories. Going outside the textual data, survey data expands new horizon of analyzing same thing from different angle, such as perceived polarization or ideological alignment.

As polarization is important to understand and extreme instances can be danger to societies, it definitely needs more limelight and it would definitely be interesting to actually carry out proper study including entire time frame and identify the change.

A Code

```
# Load packages
library(rJava)
library(mallet)
library(dplyr)
library(stringr)
library(ggplot2)
library(igraph)
library(tidyverse)
library(xtable)

# Load data
load("/20050720_congressional_records.RData")

# Drop bill label and keep only unique texts
df <- subset(records, select = -c(bill)) |> unique()

# Get text lengths
l <- str_replace_all(df$text, "[\r\n]" , "") |> str_split(" ") |> lengths()

# Look at the distribution
hist(l, breaks=200)

# Create dataframe keeping text lengths greater than 20
df <- df |> filter(l > 20)

# Tidy text by removing \n
df$text <- str_replace_all(df$text, "[\r\n]" , "")

# Stopwords
stopwords_en_file_path <- mallet_stoplist_file_path("en")

# Records instances
df.instances <-
  mallet.import(id.array = row.names(df),
                text.array = df$text,
```

```r
              stoplist = mallet_stoplist_file_path("en"),
              token.regexp = "\\p{L}[\\p{L}\\p{P}]+\\p{L}")

# Load Model
topic.model <- MalletLDA(num.topics=30, alpha.sum = 1, beta = 0.1)

# For reproducibility
topic.model$setRandomSeed(42L)

# Load instances
topic.model$loadDocuments(df.instances)

# Train
topic.model$setAlphaOptimization(20, 50)
topic.model$train(200)
topic.model$maximize(10)

# Model results
doc.topics <- mallet.doc.topics(topic.model, smoothed=TRUE, normalized=TRUE)
topic.words <- mallet.topic.words(topic.model, smoothed=TRUE, normalized=TRUE)

# Save topics and most significant words
print(xtable(mallet.topic.labels(topic.model, num.top.words = 20) %>% as_tibble(), type =
    "latex"), file = "table.tex")

# Predict the topic
df['topic'] = apply(doc.topics, 1, which.max)

# Create graph dataframe
graph_df <-
  select(df, c(speaker_bioguide, topic)) %>%
  pivot_longer(-speaker_bioguide, values_drop_na = TRUE) %>%
  pivot_wider(
    id_cols = speaker_bioguide,
    names_from = value,
    values_from = value,
    values_fn = length,
```

```
    values_fill = 0
  ) %>% as.data.frame()

# Modify rownames of dataframes
legislators <- legislators %>% as.data.frame()
rownames(legislators) <- legislators %>% as.data.frame() %>% pull(bioguide)
rownames(graph_df) <- pull(graph_df, speaker_bioguide)

# Merge legislators with graph dataframe to include all legislators
graph_df <- merge(graph_df, select(legislators, party, gender), by=0, all=TRUE)
graph_df[is.na(graph_df)] <- 0

# Bipartiate
g <- graph.incidence(subset(graph_df, select=-c(gender,party, speaker_bioguide,
    Row.names)), weighted = TRUE)

# Node colors for bipartite graph
colrs <- c("green", "cyan")[V(g)$type + 1L]

# Quick look at bipartite graph
plot(g, vertex.color = colrs,layout = layout_as_bipartite, vertex.label.cex = 0.3,
    vertex.size=2)

# Bipartiate through as_incidence_matrix
bipartite_matrix <- as_incidence_matrix(g)

# Matrix multiplaction to get projection
person_matrix_prod <- bipartite_matrix %*% t(bipartite_matrix)

# Diagonals to zero
diag(person_matrix_prod) <- 0

# Create the one-mode network of members
g_p <- graph_from_adjacency_matrix(person_matrix_prod,
                                mode = "undirected",
                                weighted = TRUE)
```

```r
# Add attributes
V(g_p)$party <- select(graph_df, party)[,]
V(g_p)$gender <- select(graph_df, gender)[,]

# Create party coloring
colrs3 <- c("blue", "red")[(V(g_p)$party == "Republican") + 1L]

# Add nice layout
lo <- layout_nicely(g_p)

# Plot the graph
plot(g_p, layout=lo, vertex.color = colrs3, vertex.label.cex = 0.3,
    vertex.size=3+1/10*degree(g_p))

# Apply cluster_edge_betweenness algorithm for modularity
ceb <- cluster_edge_betweenness(g_p)

# View created clusters
plot(ceb, g_p, layout=lo, vertex.color = colrs3, vertex.label.cex = 0.3,
    vertex.size=3+1/10*degree(g_p))

# Get modularity result
modularity(ceb)

# Get homophily results
assortativity_nominal(g_p, as.numeric(factor(V(g_p)$party)), directed=F)
assortativity_degree(g_p, directed=F)
```

B Topics

	value
1	brumidi capitol constantino work speaker brumidi's american resolution anniversary great rotunda building america florida honor gentleman ceremony life congress birth
2	court roberts president senate judge supreme john process justice nominee law years u.s judiciary nomination extraordinary american confirmation nation's important
3	water million bill world foreign international radio government safe marti darfur peace cuban programs billion operations nations khartoum health countries
4	great lakes side organizations overseas water organization democracy famous opposition focus oil drilling drinking michigan attempt taxpayer destruction october millions
5	iraq christians iraqi christian assyrian religious baghdad syria yousef church americans assyrians iraqis community rights regime saddam population communities hussein's
6	haiti senator u.n situation elections administration haitian island haiti's act bush year police problems abuses providing frankly involvement haitians view
7	guantanamo war detainees chairman bay terrorist terrorists rohrabacher california afghanistan iraq put americans terrorism law facility interrogation intelligence gentleman information
8	freedom enemies israel radical courage terrorists terror death islam serving victims kill hold bombers health training soldiers men fought honor
9	senate hearing committee july unanimous consent president national secretary authorized wednesday resources treasury meet washington office testimony a.m building session
10	state department act funding trafficking report prostitution hiv/aids sex h.r relief policy funds program reforms support legislation subcommittee work attention
11	west kosova report services pristina kosovo visa consular state international papua indonesian department papuans relations law india visas status region
12	state department government young families romania federal passport service passports certificates cases women romanian birth training order union reyes encourage
13	great lakes chavez joint florida rogers management protect freedom venezuela members network oil governors strategy civilian media federal deal information
14	china chinese states united nuclear american companies general government jobs company foreign buy zhu merger report good taiwan investment create
15	senator amendment president consent unanimous vote senators member give behalf committee amendments staff interest side work floor senate subject up-armored
16	space weapons chemical national security nuclear assets pueblo intelligence ims treaty budget depot forward outer arms protect program space-based ctbt
17	bill defense million committee military department men senate women forces program authorization year fiscal armed operations families members increase threat
18	court supreme questions person rights law decision privacy mind decisions conservative america government case civil family federal state vote lifetime
19	taylor cyprus motion table charles lay court special sierra leone agreed move human rights liberia africa crimes voted turkey west
20	iraq iraqi troops forces chairman soldiers administration security strategy iraqis success withdrawal military congress date debate american home responsibility plan
21	people support world time years country colleagues today american fact war nation issue important national hope opportunity part urge house
22	cuba justice fugitives york guard jersey crimes prison people system million watch joanne chesimard capture criminals record trade castro bring
23	great lakes water africa michigan change beautiful issues culture governors abstinence aids canada progress broad approach condoms life understand worth
24	children family students child families foster adoption senator find home money raise values care don't permanent members usaid bill parents
25	education quality care health bill act school programs college medicare system schools higher teachers provide working homeless high time work
26	iraq terrorism war fight u.s forces iraqi terrorists coalition freedom amendment commitment democracy colleagues terrorist enemy security attacks win mission
27	chairman time gentleman yield committee amendment balance consume california friend member minutes reserve colleague distinguished texas relations gentlewoman good opposition
28	indian trust bill secretary management accounts title government federal offer interest affairs accounting committee fund section individual settlement cobell indians
29	property intellectual funds countries economic foreign development lebanon reconstruction work special educational industry enforcement piracy range h.r laws developing protection
30	amendment states united u.s congress president government make authority international policy work made security foreign state important process palestinian chairman

References

Catarina Pereira, R. d. S., & Rosa, C. (2024). How to measure political polarization in text-as-data? a scoping review of computational social science approaches. *Journal of Information Technology & Politics*, *0*(0), 1–14. https://doi.org/10.1080/19331681.2024.2318404

Congress.gov. (2005). *United States Congress on July 20, 2005*. https://www.congress.gov

Curran, B., Higham, K. W., Ortiz, E., & Filho, D. V. (2017). Look who's talking: Two-mode networks as representations of a topic model of New Zealand parliamentary speeches. *PLoS ONE*, *13*. https://api.semanticscholar.org/CorpusID:23916392

Fišer, D., & Pahor de Maiti, K. (2021). Voices of the parliament: A corpus approach to parliamentary discourse research. Retrieved June 4, 2024, from https://sidih.si/20.500.12325/120

Gordon College. (n.d.). Graph theory: Basic definitions and theorems. Retrieved June 3, 2024, from https://mathworld.wolfram.com/BipartiteGraph.html

Interian, R., G. Marzo, R., Mendoza, I., & Ribeiro, C. C. (2023). Network polarization, filter bubbles, and echo chambers: An annotated review of measures and reduction methods. *International Transactions in Operational Research*, *30*(6), 3122–3158. https://doi.org/https://doi.org/10.1111/itor.13224

Lelkes, Y. (2016). Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, *80*, 392–410. https://api.semanticscholar.org/CorpusID:147357263

Maklin, C. (2022). Latent dirichlet allocation. Retrieved June 4, 2024, from https://medium.com/@corymaklin/latent-dirichlet-allocation-dfcea0b1fddc

Martino, F., & Spoto, A. (2006). Social network analysis: A brief theoretical review and further perspectives in the study of information technology. *PsychNology J.*, *4*, 53–86. https://api.semanticscholar.org/CorpusID:265814248

Murphy, P., & Knapp, B. (2018). *Bipartite/Two-Mode Networks in igraph*. Retrieved June 3, 2024, from https://rpubs.com/pjmurphy/317838

O'Malley, A. J., & Onnela, J.-P. (2014). Topics in social network analysis and network science. Retrieved June 3, 2024, from https://api.semanticscholar.org/CorpusID:18091152

Peterson, A., & Spirling, A. (2018). Classification accuracy as a substantive quantity of interest: Measuring polarization in westminster systems. *Political Analysis*, *26*, 120–128. https://doi.org/10.1017/pan.2017.39

Potgieter, D., & Cooper, V. (2022). Polarizing Information Will Be One of This Century's Great Challenges. Retrieved June 3, 2024, from https://www.templetonworldcharity.org/blog/polarizing-information-will-be-one-centurys-great-challenges

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Storopoli, J. (2019). Topic modeling: How and why to use in management research. *Revista Ibero-Americana de Estratégia*, *18*, 316–338. https://doi.org/10.5585/ijsm.v18i3.14561

Walberg, T. (n.d.). How congress works. Retrieved June 3, 2024, from https://walberg.house.gov/about/how-congress-works

Weisstein, E. W. (n.d.). Bipartite graph. Retrieved June 3, 2024, from https://mathworld.wolfram.com/BipartiteGraph.html

Zhou, T., Ren, J., Medo, M. š., & Zhang, Y.-C. (2007). Bipartite network projection and personal recommendation. *Phys. Rev. E*, *76*, 046115. https://doi.org/10.1103/PhysRevE.76.046115