RoughDraft

Arielle Dror, Kara Van Allen, Emily Daubenspeck

November 28, 2018

##Abstract

Introduction

As the year comes to a close, Washington D.C. is marking the end of the 115th Congress (2017-2018) and the halfway point of President Donald Trump's first term. There seems to be perception by the public that Congress is more partisan than ever, but is this really true?[1] Political polarization within greater society is destructive in its own capacity, but polarization within the bodies of governance tasked with serving the citizens of the United States has a number of serious implications.[2] Tremendous criticism has emerged in recent years of partisanship within Congressional bodies, due to its ability to impede the effective and efficient governing of the country as a whole.[3] However, as vitriolic rhetoric against many marginalized groups of people is endorsed by many members of one specific party, arguments have emerged that polarization is a natural and healthy response by liberal politicians to what they perceive in their colleagues to be allegiance with a party that is supportive of, or at least indifferent to, issues of racism, classism, sexism, and many other forms of insidious discrimination.[4] Inspired by Fowler's[5] work on Congressional Cosponsorship Data, we set out to determine how divided Congress really is. In order to have a manageable sample size, we use the cosponsorship data of (n bills) in the Senate during the 1115h Congress. How much cosponsorship goes on between or within parties and who seems to be facilitating this? Do Senatorial Caucuses help explain who tends to sponsor with whomst? These are the questions that we set out to answer.

Data

We obtained our data through multiple sources. All data relating to cosponsorship of bills for the 115th Congress were obtained from Legiscan.com, a website dedicated to real-time tracking of legislation moving through both houses of Congress. Through this site, we were able to download and manipulate a series of data sets that would go on to comprise our primary adjacency matrix; this series of data sets included information on each bill (name, ID, topic, committee, etc.) and cosponsor (name, ID, etc.). Joining these sets together, we produced a comprehensive list of every senator who has cosponsored a bill in the 115th Congress. We collected attributes including age, party, state represented, and the number of bills sponsored and cosponsored for each senator from Wikipedia and Congress.gov, then used this information to form our own data file and, eventually, an object containing nodal attributes for our networks.

Nodal characteristics included in our data set as as follows: State, Party, Gender, Age, Race, Number of Bills Cosponsored, Number of Bills Sponsored, and the weight of nodes within the network. The network is undirected and edge weights indicate the number of bills cosponsored mutually by pairs of senators. We intend to use Party as our most important attribute in order to assess the presence of political polarization within the Senate, but we also hope to use the attributes Gender and Race to explore the potential for uneven distribution of influence of involvement in the Senate.

[1] http://nymag.com/intelligencer/2017/05/in-the-trump-era-america-is-racing-toward-peak-polarization.html
[2] http://www.apsanet.org/portals/54/Files/Task%20Force%20Reports/Chapter2Mansbridge.pdf
[3] https://www.sciencedaily.com/releases/2015/04/150422104238.htm
[4] https://www.vox.com/polyarchy/2016/3/24/11298808/american-politics-peak-polarization
[5] http://fowler.ucsd.edu/cgnet.pdf

Methods

Completed already

Currently, we have created our weighted adjacency matrix and one visualization. While the visualization is important to have, it has been challenging to create one which is easy to understand and interpret. The very dense nature of our networks creates a unique visualization challenge which we are working to overcome. It is our intention to next try to visualize parts of the network, and narrow down the number of interactions we are graphing in our model, in an attempt to decrease the density of the plot. By creating a cutoff of minimum number of bills co-sponsored between two senators, we will be able to decrease the density like we want to. We are currently trying to figure out how to determine that cutoff to stay true to the integrity of the data and our question. A majority of the edge weights (number of co-sponsored bills) are between 0 and 17, while the largest one is above 200. This large range is why we are trying to narrow the number of different edge weights for the visualization.

What we'd like to do moving forward

Partisanship is inherently a question of whether observable communities or groups within a greater network can be readily identified, and whether those groups are divisible along party lines. Therefore, our focus going forward will be first and foremost on constructing a workable SBM model, which will allow us to observe and interpret a model grounded on the idea that group divisions are a defining feature of social network models.

After we achieve an interpret able SBM and are able to evaluate the information that it indicates, we may attempt two additional methods for modeling. First, an ERGM specifically exploring the likelihood of cosponsership between members of the same party (i.e. nodematch(party), among other terms) will allow us to further investigate the role of a variety of nodal characteristics on connections within the cosponsership network. Additional ERGM terms will include ctriple (to evaluate clustering and presence of cyclic triples in the network), degree (with by= and homophily=TRUE specified to examine degree distributions within subnetworks of nodes bearing particular characteristics), and nodematch (for party, race, gender, and age), to test for the significance of various nodal attributes on connection formation, especially those attributes heretofore undiscussed in papers on this subject (i.e. gender and race).

To construct these models will be all well and good, but if we seek to discuss network topology in any kind of meaningful way, we have determined that it will be necessary to take a sample of our total set of cosponsered bills and proceed to visualize them. These samples will likely incorporate something close to 10% of the total bill volume in our current adjacency matrix (10% is generally recognized as the level at which a sample is able to be generalized as a representation of a greater population). Visualizing a series (~3) of these networks will allow us to use algorithmic approaches to detect the presence of communities within them and to make comments on the impact that various nodal characteristics appear to have on community formation as a whole.

Results

The way that we approached this project was fairly inefficient, where data collection is concerned. First and foremost, we were inspired by a paper by James Fowler on social network analysis of groups in Congress, and anticipated that the data he had used, or the repositories that had previously provided him access to the necessary information, would be updated to include bills from the 115th Congress, as it was our goal to incorporate Fowler's methods into a study of the most recent data available. Unfortunately, it was extremely difficult to find the necessary data in accessible, free, or otherwise reasonably usable formats, and this ate up a gigantic chunk of our time during the beginning of this project. As such, production of actual results has been slow going. Now that we have the data and have formulated it into a proper adjacency matrix, we anticipate that progress will increase at a steady pace. However, for the time being, we will be focusing our energy most closely on refining our methodological approaches and will report on any relevant results at a later time.

% Table created by stargazer v.5.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Mon, Dec 17, 2018 - 1:25:39 PM

% Table created by stargazer v.5.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Mon, Dec 17, 2018 - 1:27:01 PM

% Table created by stargazer v.5.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Mon, Dec 17, 2018 - 1:27:21 PM

Discussion

–Stay tuned!–

Citations

Table 1: Limtited Highly-weighted Model

	Term coefficients:
	top_net
nodefactor.Gender.M	-0.301^{***}
	(0.081)
nodefactor.Race.Asian	-1.528^{***}
	(0.236)
nodefactor.Race.Hispanic	0.526^{***}
	(0.194)
nodefactor.Race.White	-0.114
	(0.361)
nodematch.Gender	0.053
	(0.096)
nodematch.Race	-0.084
	(0.359)
absdiff.Age	-0.010^{**}
	(0.004)
mix.Party.Democratic.Democratic	0.483
	(0.479)
mix.Party.Democratic.Independent	1.090^{*}
	(0.583)
mix.Party.Independent.Independent	10.617
mix.Party.Democratic.Republican	-1.913^{***}
	(0.476)
mix.Party.Independent.Republican	-2.497^{***}
	(0.581)
mix.Party.Republican.Republican	-1.499^{***}
	(0.483)
nodecov.Age	0.011^{***}
	(0.002)
isolates	6.589
Akaike Inf. Crit.	5,646.342
Bayesian Inf. Crit.	5,745.420
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 2: Complex Highly-weighted Model

	Term coefficients:
	top_net
nodefactor.Gender.M	-0.254^{**}
	(0.109)
nodefactor.Race.Asian	-1.094^{***}
	(0.390)
nodefactor.Race.Hispanic	0.398
	(0.248)
nodefactor.Race.White	-0.026
	(0.364)
nodematch.Gender	0.032
	(0.111)
nodematch.Race	-0.093
	(0.348)
absdiff.Age	-0.009^{**}
	(0.004)
mix.Party.Democratic.Democratic	-0.186
	(0.562)
mix.Party.Democratic.Independent	0.236
	(0.675)
mix.Party.Independent.Independent	0.739
	(0.893)
mix.Party.Democratic.Republican	-1.589^{***}
	(0.535)
mix.Party.Independent.Republican	-2.169^{***}
	(0.648)
mix.Party.Republican.Republican	-1.005^{*}
	(0.553)
nodecov.Age	0.011^{***}
	(0.003)
isolates	824.160
gwesp	-0.017^{***}
	(0.005)
gwesp.decay	8.433
	(38.097)
Akaike Inf. Crit.	5,756.546
Bayesian Inf. Crit.	5,868.834
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 3: Original Highly-weighted Model

	Term coefficients:
	top_net
nodefactor.Gender.M	-0.299^{***}
	(0.081)
nodefactor.Race.Asian	-1.506^{***}
	(0.234)
nodefactor.Race.Hispanic	0.531^{***}
	(0.189)
nodefactor.Race.White	-0.088
	(0.359)
nodematch.Gender	0.048
	(0.095)
nodematch.Race	-0.104
	(0.362)
absdiff.Age	-0.010^{***}
	(0.004)
mix.Party.Democratic.Democratic	0.478
	(0.480)
mix.Party.Democratic.Independent	1.067^{*}
	(0.568)
mix.Party.Independent.Independent	10.828
	(196.968)
mix.Party.Democratic.Republican	-1.919^{***}
	(0.482)
mix.Party.Independent.Republican	-2.490^{***}
	(0.577)
mix.Party.Republican.Republican	-1.507^{***}
	(0.487)
nodecov.Age	0.011^{***}
	(0.002)
Akaike Inf. Crit.	5,784.216
Bayesian Inf. Crit.	5,876.689
Note:	*p<0.1; **p<0.05; ***p<0.01