

CongressRank: A PageRank Application

Chandler Chen, David Tseng

March 15, 2017

Introduction

When a bill is introduced to Congress, the original proponent is usually called the *sponsor*. Congress members who support the bill can sign onto the bill as a *cosponsor*. The political community places great emphasis on cosponsorships, as it is believed that a bill with more cosponsorships demonstrates a greater support base. Thus, the number of cosponsorships of a bill is an indicator of the popularity of the bill among Congress members. Since a Congress member will only agree to cosponsor a sponsor's bill if he or she agrees with the bill's contents, a cosponsorship can also symbolize the cosponsor's support for and agreement with the sponsor. We believe that by analyzing the network of cosponsorships, we can have a better understanding of the power dynamics and unity within Congress.

In this project, we used PageRank to analyze the cosponsorship graph. PageRank is a ranking algorithm developed to find the stationary distribution of a Markov Chain, where a crawler visits other nodes (where neighboring nodes are chosen uniformly at random) for many iterations. As the number of iterations increases, the percentage of visits a node receives converges to the graph's stationary distribution, and nodes are ranked based on the number of visits they receive. We believe that the rankings generated by PageRank on this graph will be able to rank each member's importance/support within Congress.

Methods

In our project, we represented each Congress member as a node. If member u cosponsors member v on a bill, then a directed edge is placed from u to v on the graph. If u cosponsors v on x different bills, there will be x parallel directed edges from u to v . Since outgoing edges are chosen uniformly at random in PageRank, if there are more parallel directed edges from u to v , the crawler will have a higher probability of visiting v from u . Representatives and Senators are analyzed in separate graphs.

We decided to use data from the 114th Congress (2015-2017), since this was the most recent Congress that completed its term. Although the 114th Congress had 100 Senators, 435 Representatives, and 6 non-voting members, there were 7 additional members who left

Congress before their term ended. Therefore, we included a total of 548 members (nodes) in our graph. To extract the bill and member data, we used ProPublica’s Congress API, which presented us the data in a convenient JSON format.

Pseudocode

```

1: procedure CREATEGRAPH(bills)
2:   graph = dictionary initialized to a mapping between each member to an empty
      list. This will be a mapping between each member to a list of members that he/she
      cosponsored, i.e., a dictionary mapping each node to a list of adjacent nodes.
3:   for bill in bills do
4:     for cosponsor in bill[cosponsors] do
5:       graph[cosponsor].append(sponsor)
6:   return graph
7:
8: procedure PAGERANK(bills, numIterations, p) // bills = list of bills that we are
      analyzing for our graph. numIterations = total number of visits to be made. p =
      probability of resetting to a random node.
9:   graph = createGraph(bills)
10:  membersNumVisits = dictionary initialized to a mapping between each member to
      0. This will eventually be a dictionary mapping each member to the number of visits
      he/she receives during the PageRank algorithm.
11:  currentMember = randomly chosen member
12:  membersNumVisits[currentMember] += 1
13:  for i in range(numIterations) do
14:    if currentMember sponsored no one or Uniform[0,1] < p then
15:      currentMember = randomly chosen member
16:    else
17:      neighbors = graph[currentMember] // list of adjacent nodes
18:      currentMember = randomly chosen member from neighbors, uniformly at
      random
19:      membersNumVisits[currentMember] += 1
20:  Sort members based on the number of visits, in descending order.
21:  return membersNumVisits

```

Above is the pseudocode we used to generate our graph and run our PageRank algorithm. We retrieved the list of bills and their sponsors and cosponsors by using calls to ProPublica’s Congress API.

Experiment 1: Results and Analysis

For Experiment 1, we included cosponsorships of all bills that were introduced by the 114th Congress, and all of their corresponding sponsors and cosponsors. We set $p=0.15$ (probability of resetting to a random node) and ran the PageRank algorithm for 100,000 steps. If we decrease p to 0.05 or increase p to 0.20, the rankings remain relatively the same. If we further increase p to even higher values, the results become more uniformly random, so we believe it is more meaningful to keep p at around 0.15.

| Rank | Normalized No. Visits | Name | Party |
|------|-----------------------|----------------------|-------|
| 1 | 1.000000 | Diane Black | R |
| 2 | 0.812907 | Erik Paulsen | R |
| 3 | 0.758962 | Sam Johnson | R |
| 4 | 0.744396 | Charles Boustany Jr. | R |
| 5 | 0.739433 | Brett Guthrie | R |
| 6 | 0.690203 | Christopher Smith | R |
| 7 | 0.664008 | Tom Price | R |
| 8 | 0.657639 | Paul Gosar | R |
| 9 | 0.638455 | Kevin Brady | R |
| 10 | 0.625765 | Peter Roskam | R |

Table 1: Top 10 ranking of Representatives, using all bills that were introduced

| Rank | Normalized No. Visits | Name | Party |
|------|-----------------------|------------------|-------|
| 1 | 1.000000 | Orrin Hatch | R |
| 2 | 0.983083 | Charles Grassley | R |
| 3 | 0.800678 | John Thune | R |
| 4 | 0.696344 | Benjamin Cardin | D |
| 5 | 0.693545 | Mark Kirk | R |
| 6 | 0.652788 | John Cornyn | R |
| 7 | 0.625090 | Marco Rubio | R |
| 8 | 0.620461 | Jerry Moran | R |
| 9 | 0.614111 | Mike Lee | R |
| 10 | 0.608837 | Rob Portman | R |

Table 2: Top 10 ranking of Senators, using all bills that were introduced

Tables 1 and 2 list the resulting top 10 Representatives and Senators from our data. The second column displays the number of visits each member received during the course of the PageRank algorithm, normalized by the maximum number received. From our results, it appears that Diane Black and Orrin Hatch are the most popular in their respective cham-

bers. One interesting thing to note is that almost every top 10 member is a Republican, even though we included members from both parties. Since we analyzed all bills that were introduced, this captures the most interactions possible between members via bills. From this analysis, it seems like Republicans were generally more unified compared to Democrats during the 2015-2017 Congress. This makes sense, because Republicans controlled both the House and the Senate during this period of time. Politicians tend to support other politicians in the same party, so therefore the Republicans’ advantage in numbers boosted their rankings.

We also noted that some members with leadership positions (such as Nancy Pelosi, Paul Ryan, Mitch McConnell) actually had relatively low rankings. For example, Mitch McConnell ranked 91 within the 100 Senators. It seems like ranking members based on a network of cosponsorships of all bills is not necessarily a good indication of actual influence in Congress. This may be because we included every bill that is introduced, so a person who writes many insignificant bills can accumulate more cosponsorships, leading to skewed results.

Experiment 2: Results and Analysis

For our next experiment, we decided to only analyze cosponsorships of bills that were actually signed into law. We ran PageRank on this filtered dataset, using the same parameters as before.

| Rank | Normalized No. Visits | Name | Party |
|------|-----------------------|-------------------|-------|
| 1 | 1.000000 | Sam Johnson | R |
| 2 | 0.801382 | Bill Posey | R |
| 3 | 0.741576 | Carolyn Maloney | D |
| 4 | 0.715829 | Terri Sewell | D |
| 5 | 0.706899 | Jeff Fortenberry | R |
| 6 | 0.590983 | Brett Guthrie | R |
| 7 | 0.567227 | Martha McSally | R |
| 8 | 0.548026 | Seth Moulton | D |
| 9 | 0.507255 | Edward Royce | R |
| 10 | 0.337657 | Christopher Smith | R |

Table 3: Top 10 ranking of Representatives, using only bills that were signed into law.

Tables 3 and 4 show the results of this experiment. It now appears that Sam Johnson and John Cornyn have the highest ranking in their respective chambers. We note that the proportion of Democrats in the top 10 category is now higher than it was in the previous experiment. One possible explanation is that by only considering cosponsorships on bills that were eventually signed into law, the rankings are more representative of the members’

| Rank | Normalized No. Visits | Name | Party |
|------|-----------------------|--------------------|-------|
| 1 | 1.000000 | John Cornyn | R |
| 2 | 0.713656 | Roy Blunt | R |
| 3 | 0.683725 | Orrin Hatch | R |
| 4 | 0.627329 | Dianne Feinstein | D |
| 5 | 0.525480 | Mazie Hirono | D |
| 6 | 0.492836 | Benjamin Cardin | D |
| 7 | 0.421928 | Sheldon Whitehouse | D |
| 8 | 0.320936 | Patrick Leahy | D |
| 9 | 0.280296 | Heidi Heitkamp | D |
| 10 | 0.278308 | Rob Portman | R |

Table 4: Top 10 ranking of Representatives, using only bills that were signed into law.

level of influence than their popularity. This idea is also supported by the fact that people in leadership positions (Paul Ryan, Nancy Pelosi, Mitch McConnell) all have higher rankings here compared to the previous experiment. For example, while Mitch McConnell was ranked 91st previously, he now ranks 15th. This shows that indeed, the second experiment shows more about each member’s influence in Congress. Because the Democrats make up approximately half of the people in the top 10, it seems like the Democrats have almost equal influence as Republicans.

Discussion and Limitations

There are several limitations to our approaches. One limitation is that although we were only aiming to analyze the 114th Congress, many members were already in Congress before 2015, and they had many more contributions before then. So while our rankings may show some general trends in popularity/support from the interactions in 2015-2017, it does not include interactions between members that were already there before 2015.

Our model is also limited by omitting the relative importance of each considered bill. For instance, H.R.6135-114, a bill *To designate the Federal building and United States courthouse located at 719 Church Street in Nashville, Tennessee, as the "Fred D. Thompson Federal Building and United States Courthouse"* may not be as important as a major defense bill. True, this notion of importance is subjective, so objective importance may be better approximated with the quantifiable amount of discretionary spending put forth in each bill. Even this approach is flawed since large trade deals or important foreign policies may be excluded, but it remains a step in the right direction.

Overall, despite these limitations, our PageRank algorithm produced some very interesting results. While Republicans may have enjoyed higher levels of popularity in the 114th Congress, Democrats and Republicans shared somewhat equal levels of influence.