CongressRank: A PageRank Application

Chandler Chen, David Tseng

March 14, 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 popularity/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 distinct directed edges from u to v. 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

```
procedure createGraph(bills)
    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.
    for bill in bills do
        u = bill[sponsor]
        for cosponsor in bill[cosponsors] do
            graph[cosponsor].append(sponsor)

procedure PageRank(bills, numIterations, p) // bills = list of bills that we are analyzing for our graph. numIterations = total number of visits to be made.
    graph = createGraph(bills). p = probability of resetting to a random node
    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.
    currentMember = randomly chosen member
    membersNumVisits[currentMember] += 1
    for i in range(numIterations) do
        if ( thencurrentMember sponsored no one or random() < p)
            currentMember = randomly chosen member
        else
            neighbors = graph[currentMember] // list of adjacent nodes
            currentMember = randomly chosen member from neighbors
        memberNumVisits[currentMember] += 1
    Sort members based on the number of visits, in descending order.
```

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 all bills that were introduced by the 114th Congress, and all of their corresponding sponsors and cosponsors. We used a p=0.15 (probability of resetting to a random node) and ran the PageRank algorithm for 100,000 steps. If we decrease p 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 the p value to 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. 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 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 bills that were actually signed into law.

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.