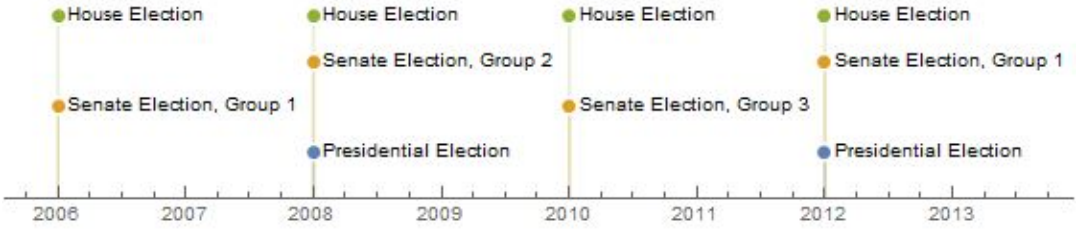
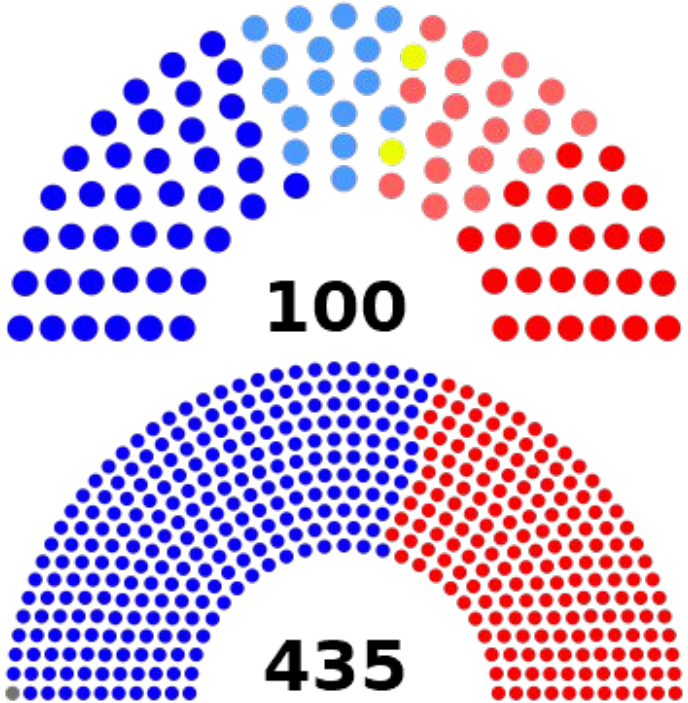
Investigation of US Political Donation Data

Leah Balter, William Cortes, Jing Lu,
Aviva Prins, and Oladimeji Salako

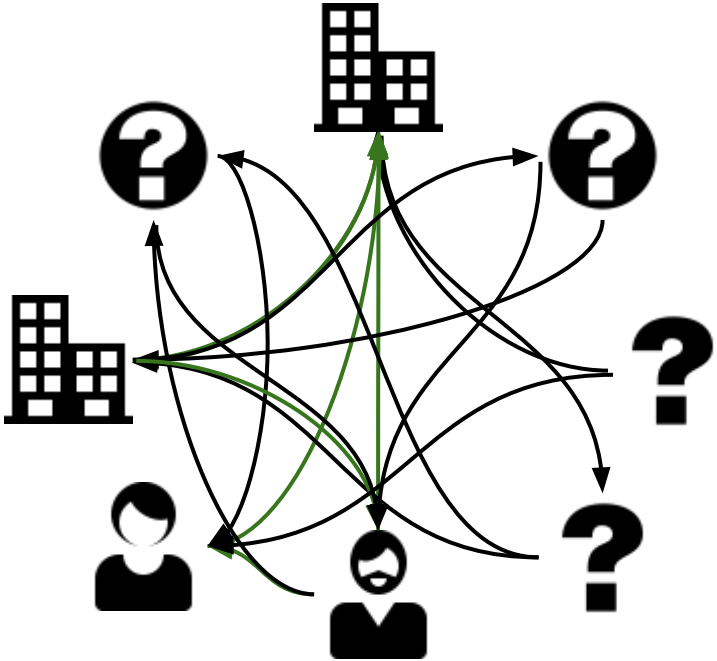
Background


House Election
Senate Election, Group 1
Presidential Election
House Election
Senate Election, Group 2
Presidential Election
House Election
Senate Election, Group 3
House Election
Senate Election, Group 1
Presidential Election
2006
2007
2008
2009
2010
2011
2012
2013

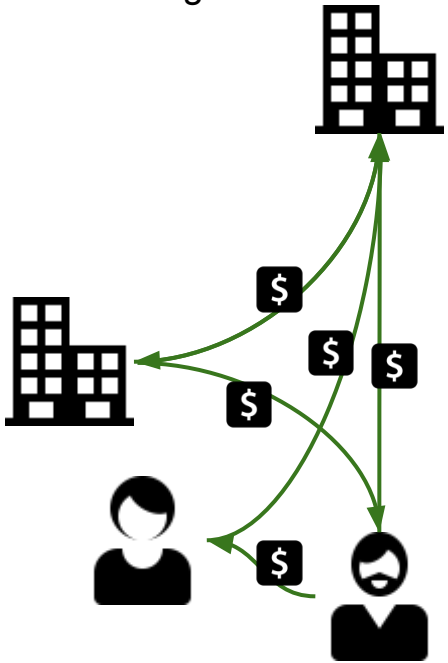

100
435

The Dataset

Federal Election Commission Data

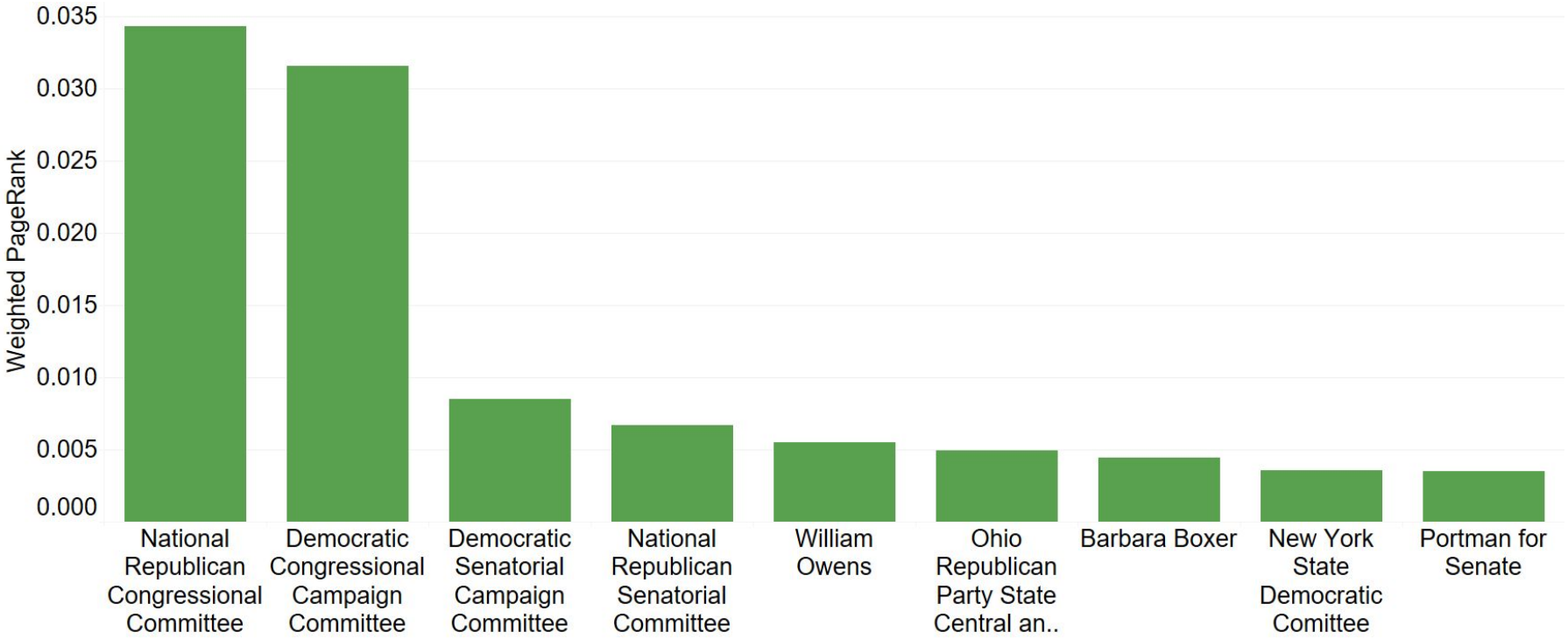


Directed Weighted Network



simpleicon.com

Weighted PageRank



Community Detection

Walktrap Algorithm

- Applicable for directed weighted networks
- Efficient for complex networks
- Maximizes modularity score
- Captures information on the community structure
- The algorithm is based on a random walk

Walktrap Computation

Initialization We start form a partition $\mathcal{P}_1 = \{\{v\} \mid v \in V\}$ of network with n communities and each community has a single vertex. t, a given constant denoting the number of steps in random walk.

For $k = 1, \cdots, n-1$ **iterations**

- Compute the distance between all adjacent vertices using $r_{ij}(k) = \sqrt{\sum_{k=1}^{n} \frac{\left(p_{ik}^t - p_{jk}^t\right)^2}{\deg(k)}}$
- Choose two communities C_1 and C_2 in $\mathcal{P}_k$ and compute the distance between communities

$$r_{C_1,C_2} = \sqrt{\frac{\sum_{k=1}^{n}\left(p_{C_1,k}^t - p_{C_2,k}^t\right)^2}{\deg(k)}},$$

  where p_{ij}^t is the probability of moving from node i to node j in t steps.
- Merge these two communities into a new community $C_3 = C_1 \cup C_2$ by minimizing the mean σ_k of the squared distance

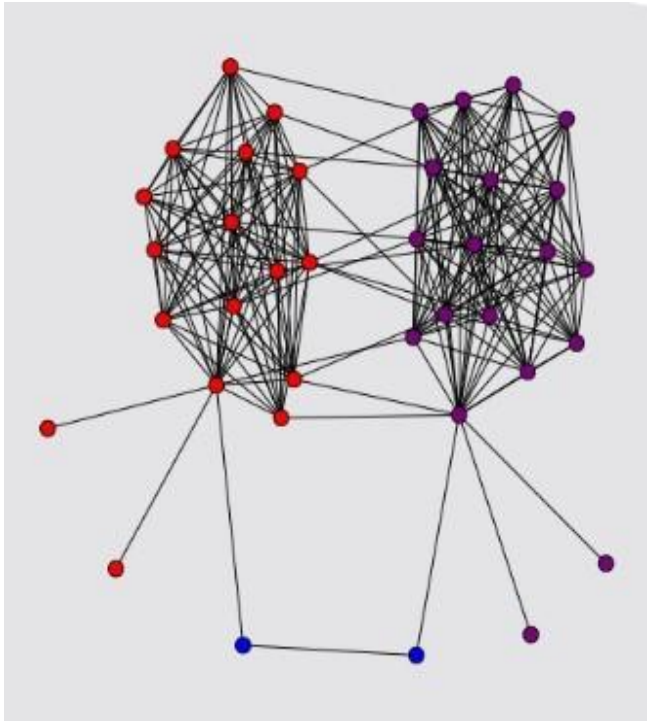
$$\sigma_k = \frac{1}{n} \sum_{C \in \mathcal{P}_k} \sum_{j \in C} r_{j,C}^2$$

  Note: it's a greedy algorithm since it tries to solve the problem of maximizing σ_k for each k
- After merging, it creates a new partition $\mathcal{P}_{k+1} = \left(\mathcal{P}_k \backslash \{C_1, C_2\}\right) \cup \{C_3\}$

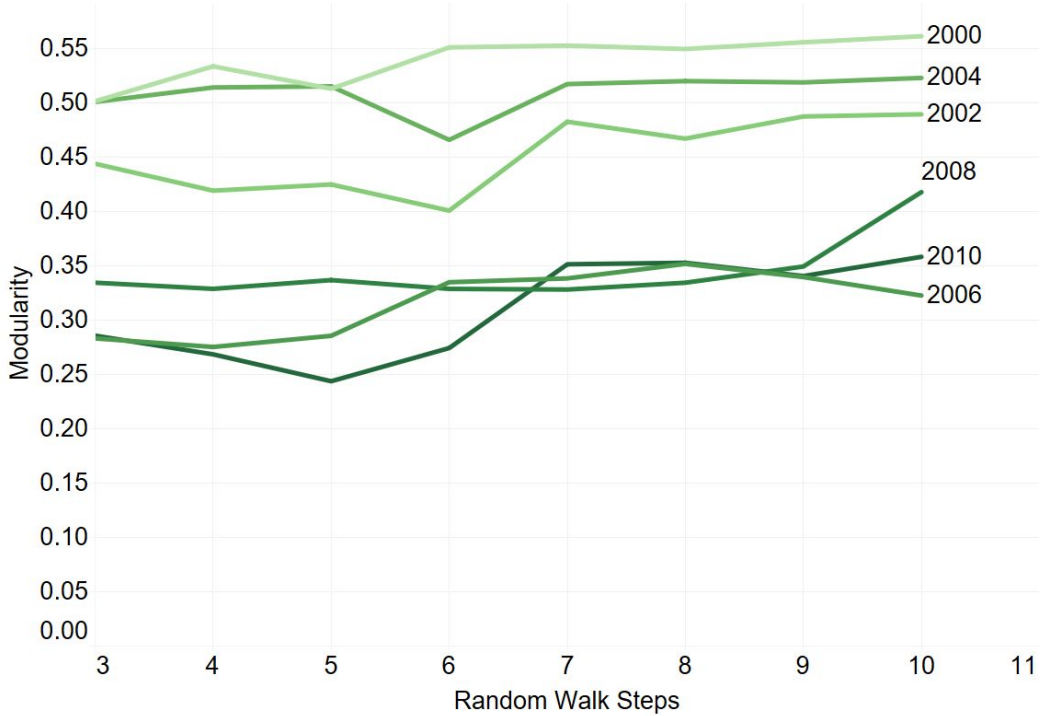
$$\triangle\sigma(C_1, C_2) = \frac{1}{n}\left(\sum_{i \in C_3} r_{j,C_3}^2 - \sum r_{i,C_1}^2 - \sum_{i \in C_2} r_{i,C_2}^2\right)$$

- Update the distance between communities.

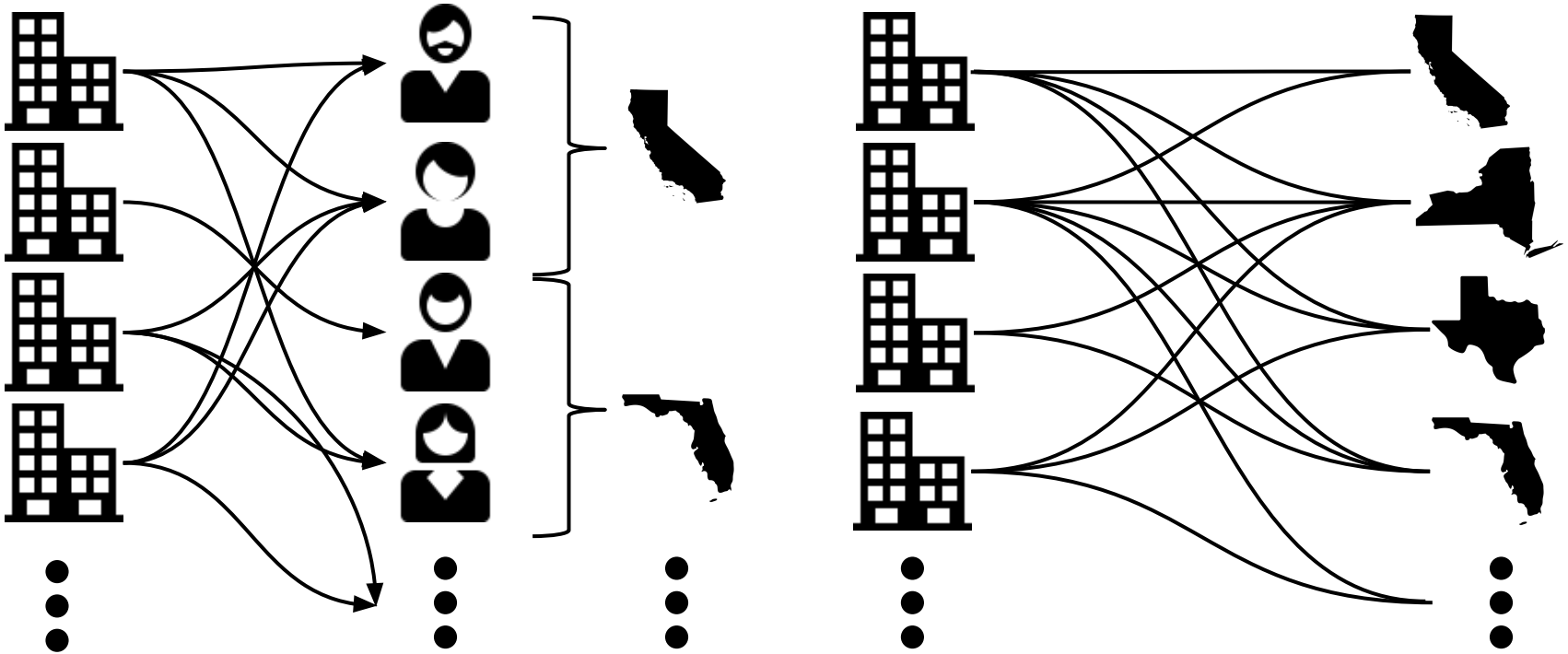
Output $\mathcal{P}_n$.



Walktrap Algorithm Result



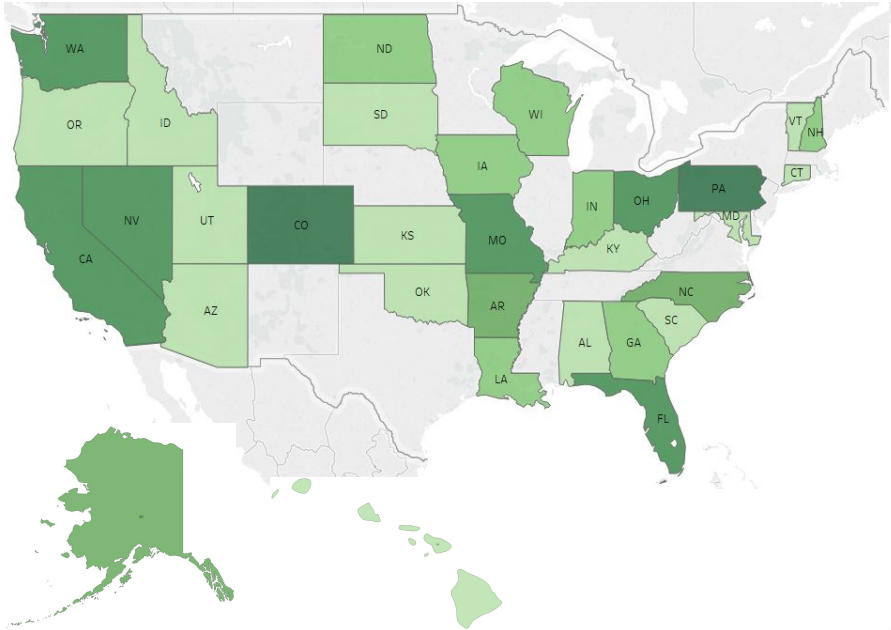
The Senate Network



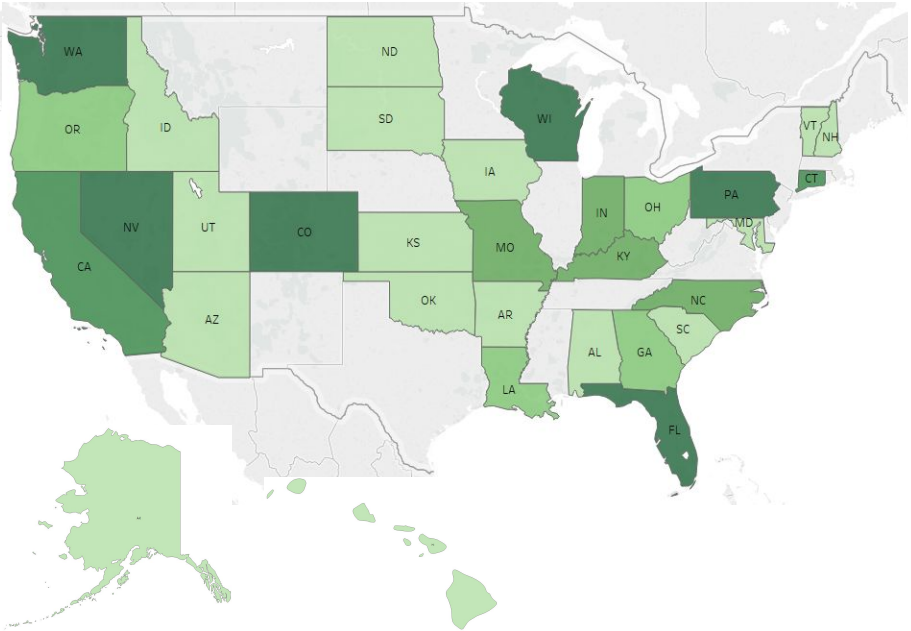
simpleicon.com
thenounproject.com

Close Elections

Weighted PageRank (2010)

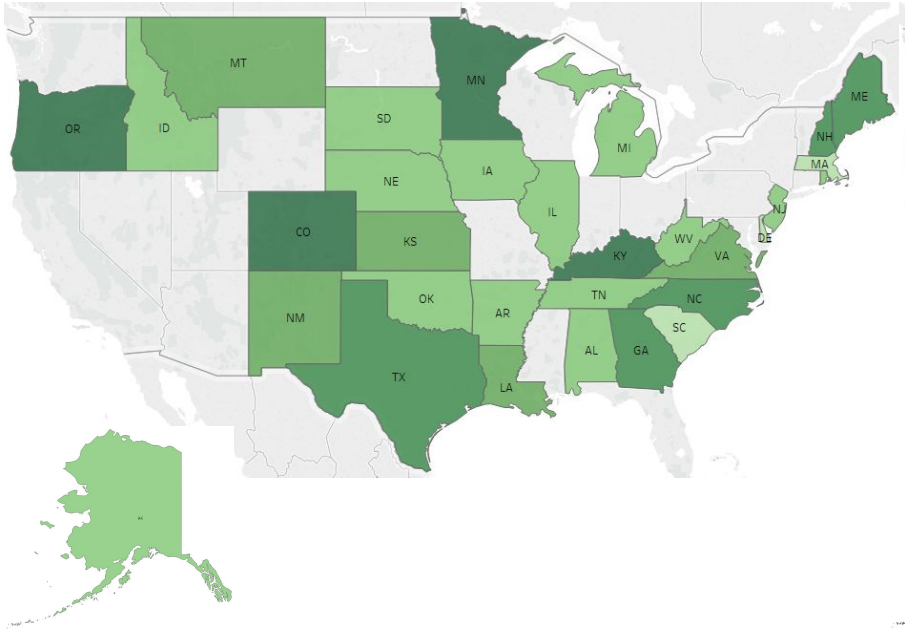


Winning Margin (2010)

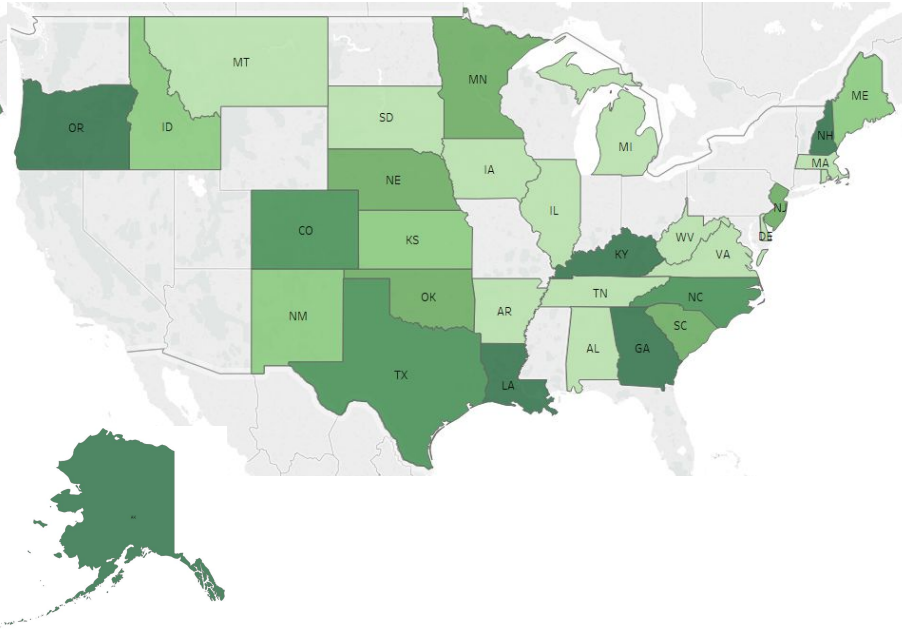


Close Elections

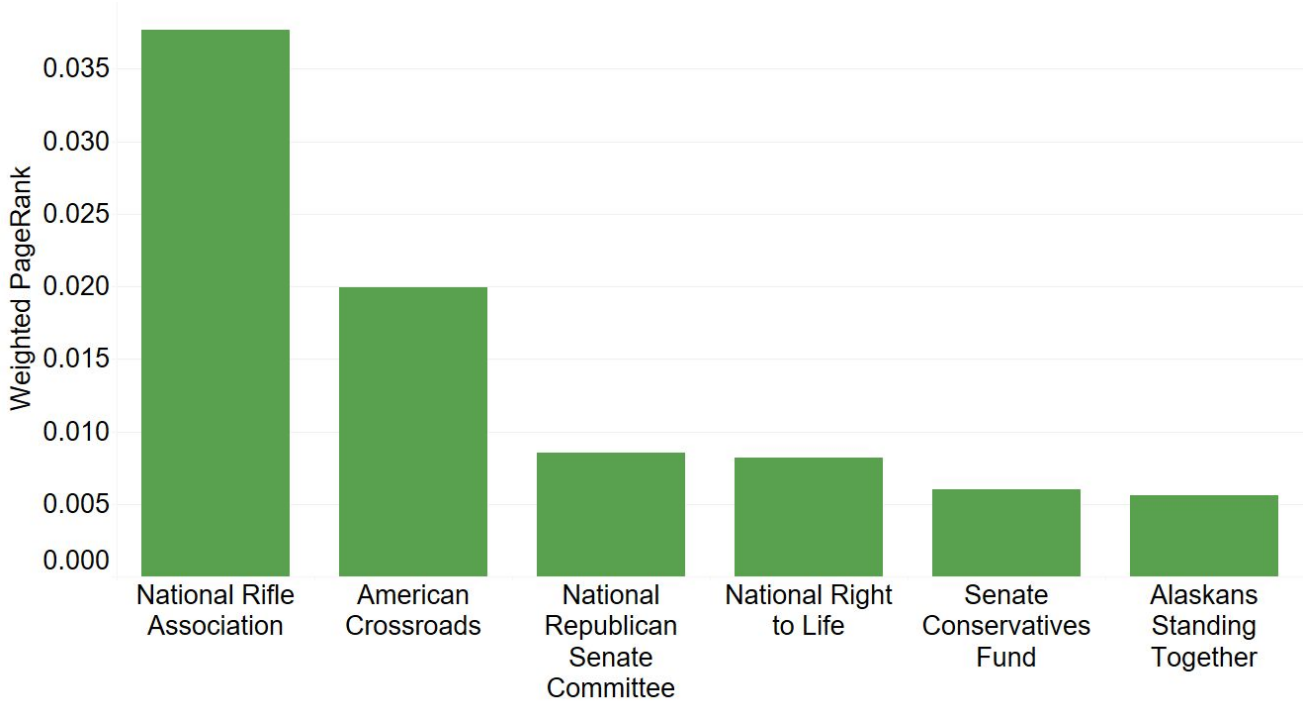
Weighted PageRank (2008)



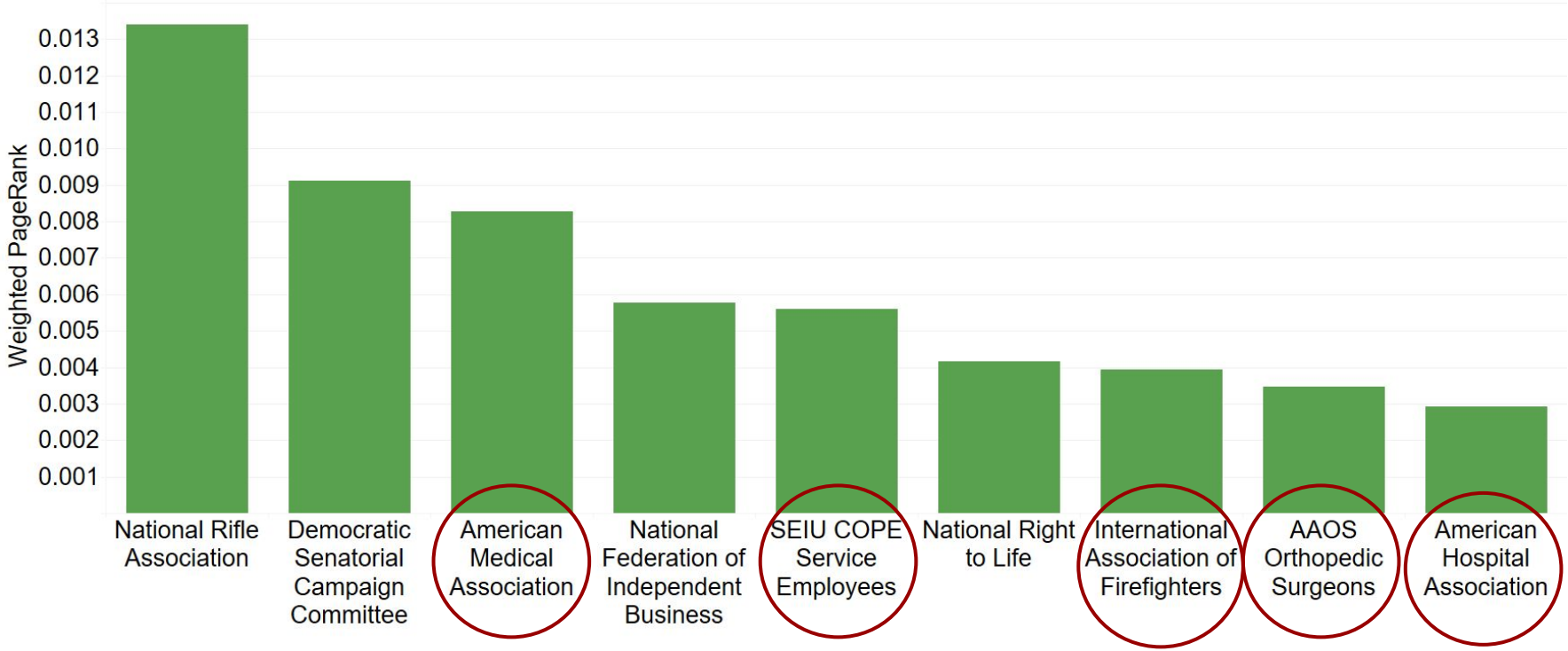
Winning Margin (2008)



Important PACs by Weighted PageRank (2010)

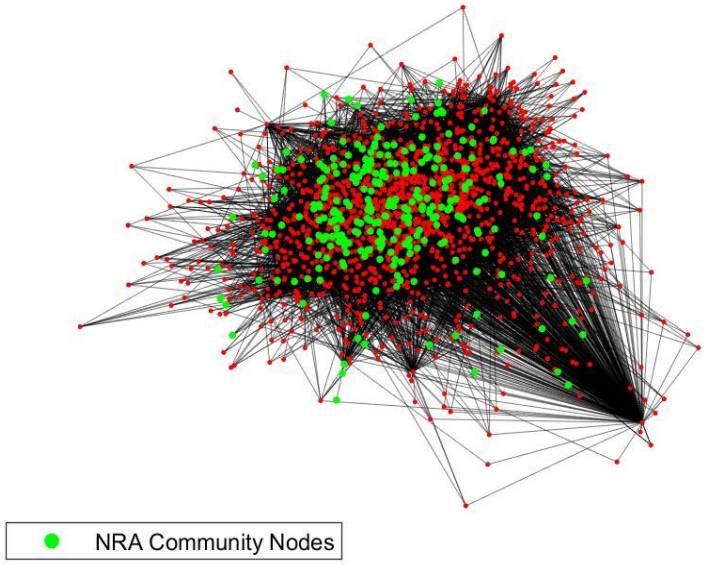


Important PACs by Weighted PageRank (2008)

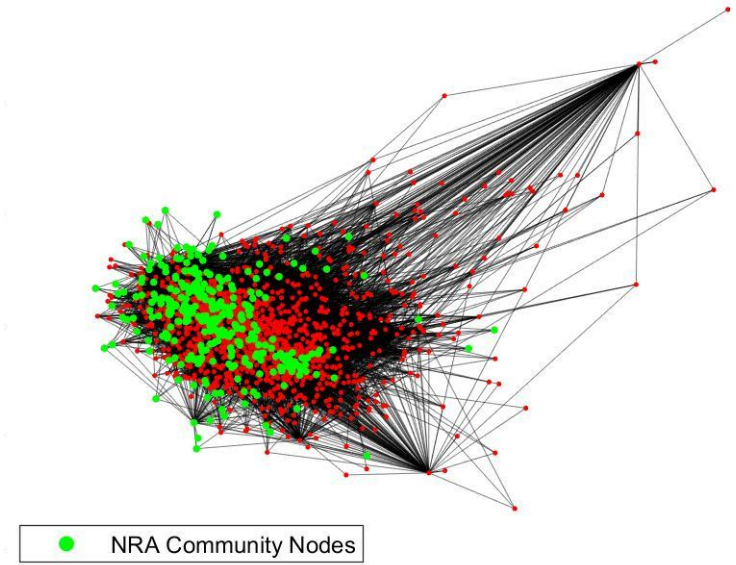


NRA Louvain Communities

2010

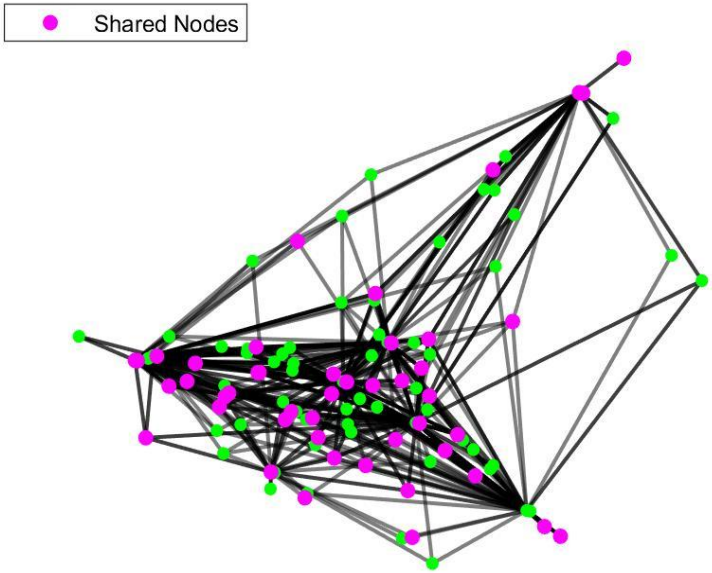


2008

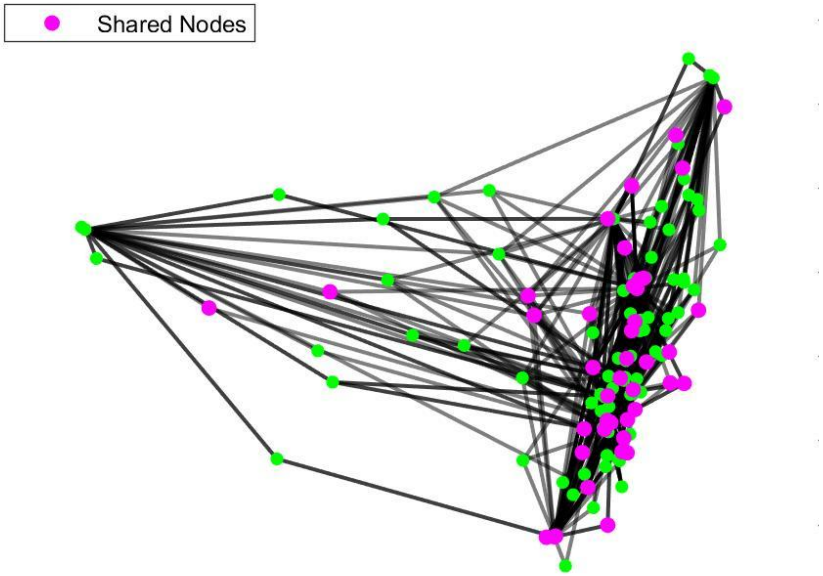


Shared NRA Community

2010

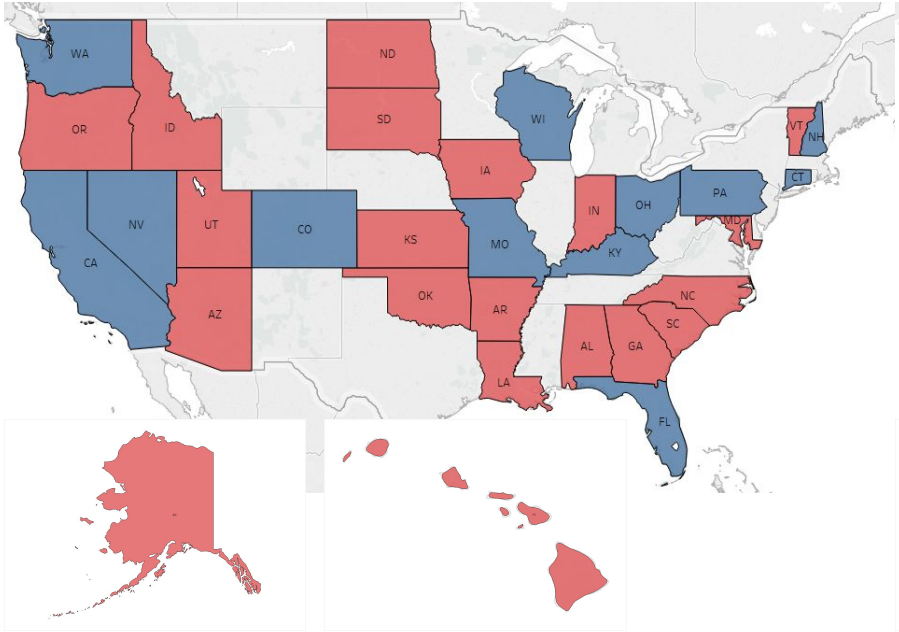


2008

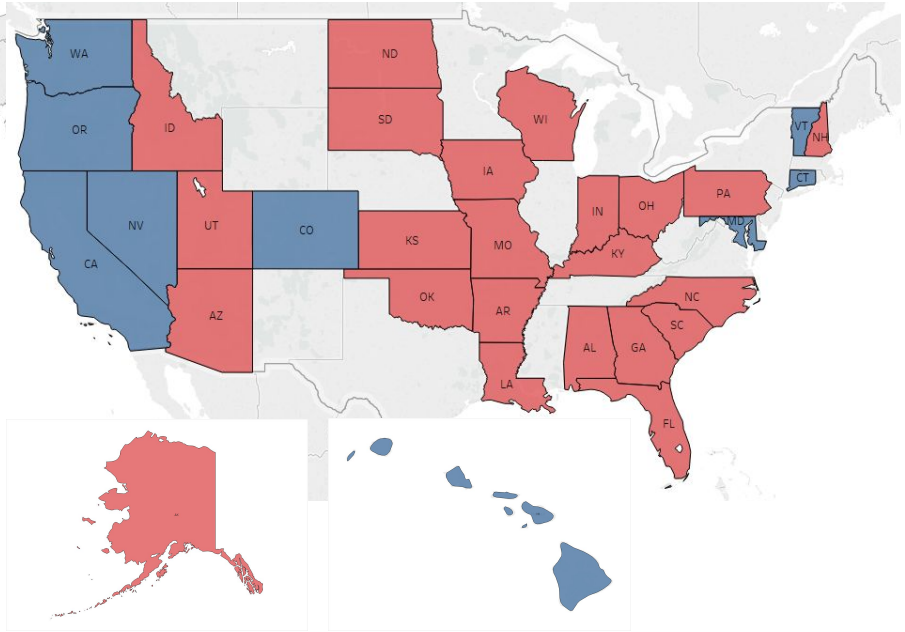


Community Detection on States

Modularity (2010)

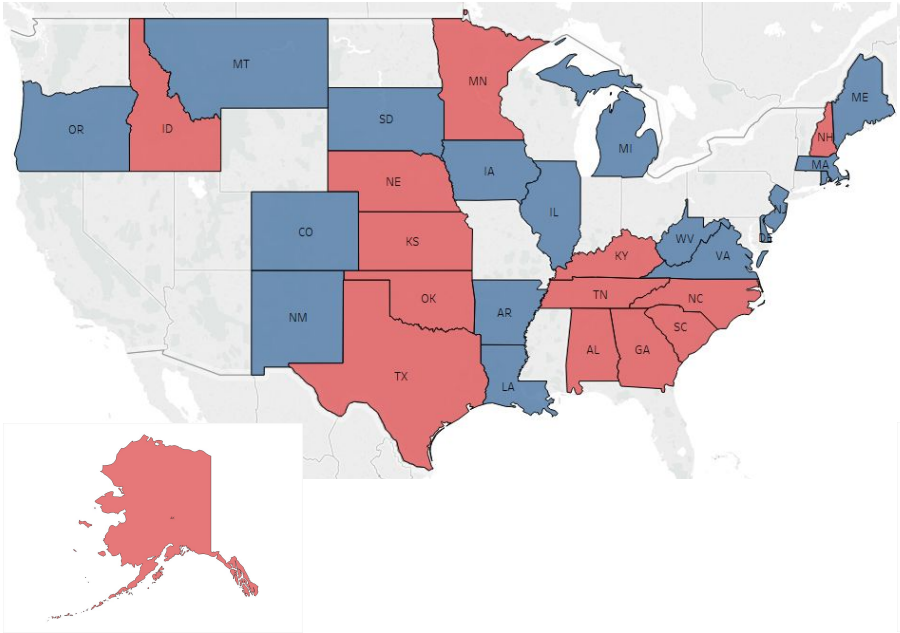


Election Result (2010)

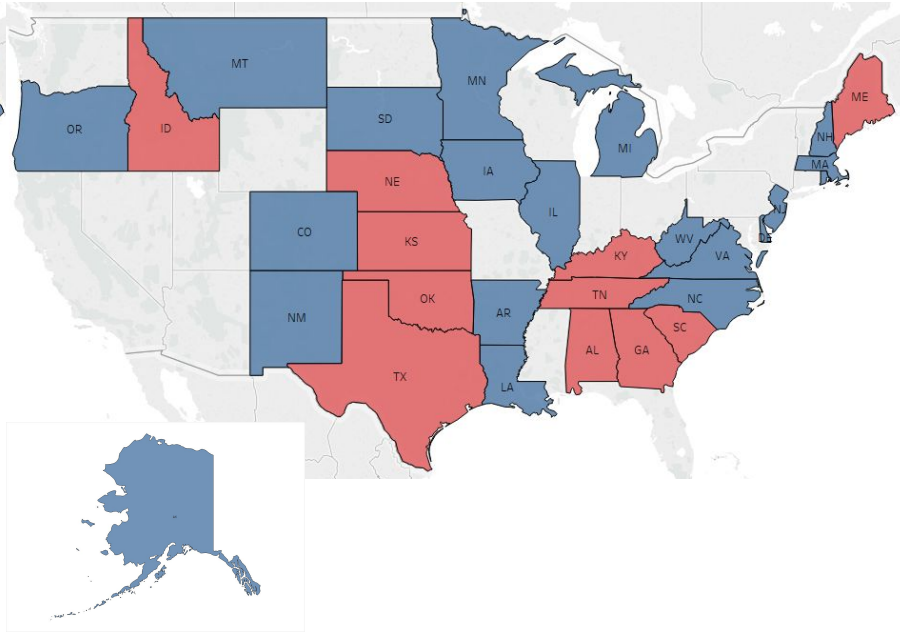


Community Detection on States

Modularity (2008)

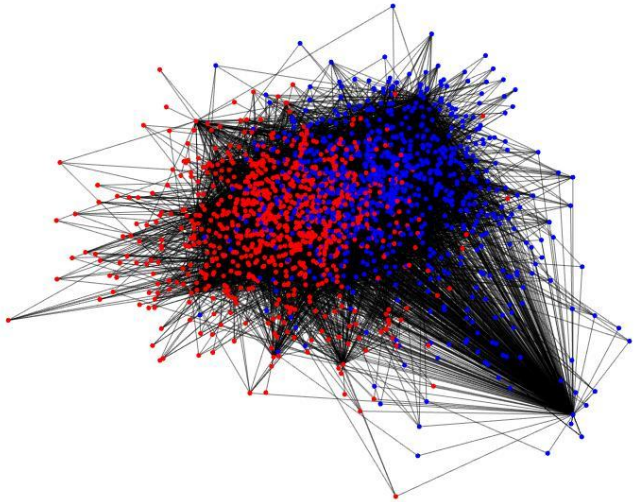


Election Result (2008)

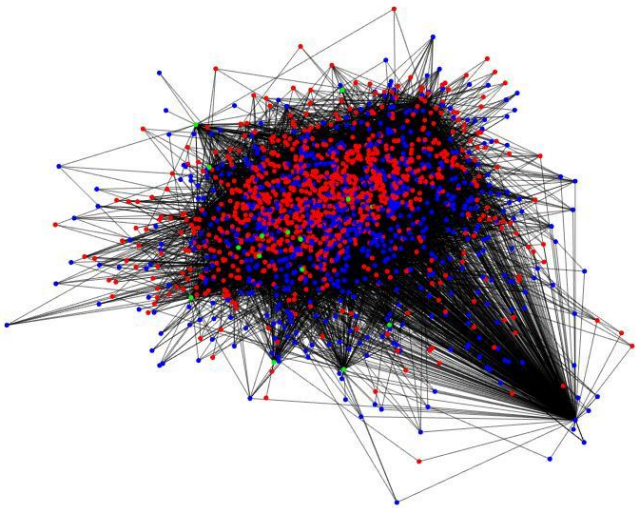


Community Detection on States and Committees

Modularity (2010)

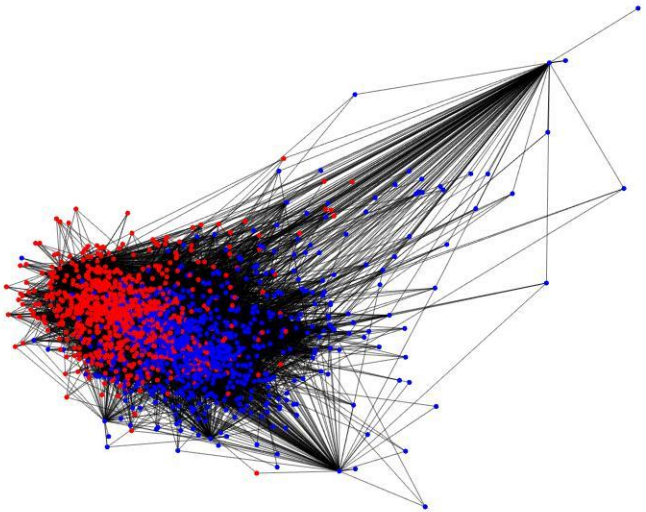


Party Affiliation (2010)

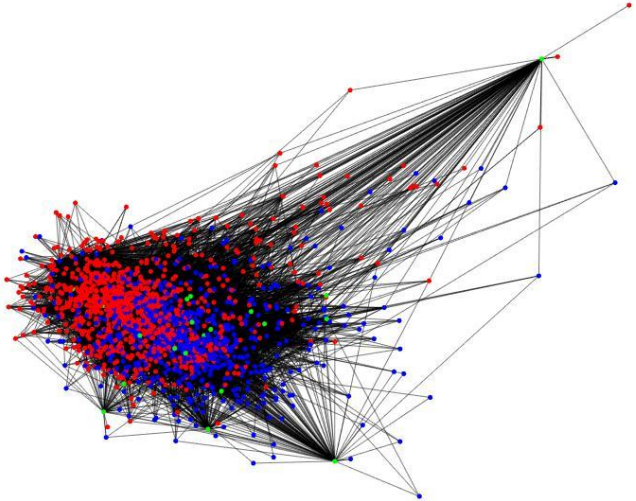


Community Detection on States and Committees

Undirected Modularity (2008)

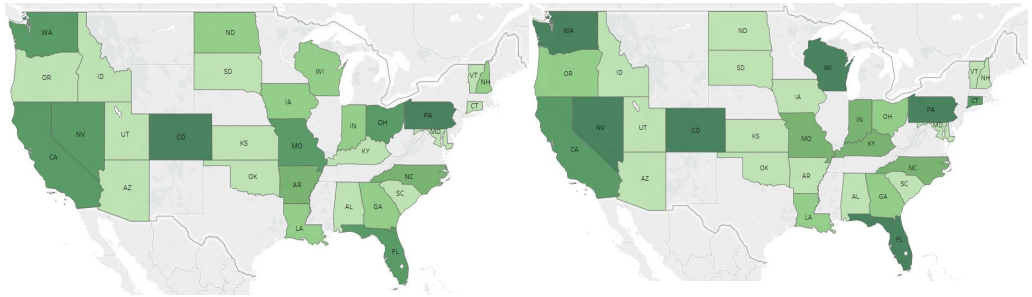


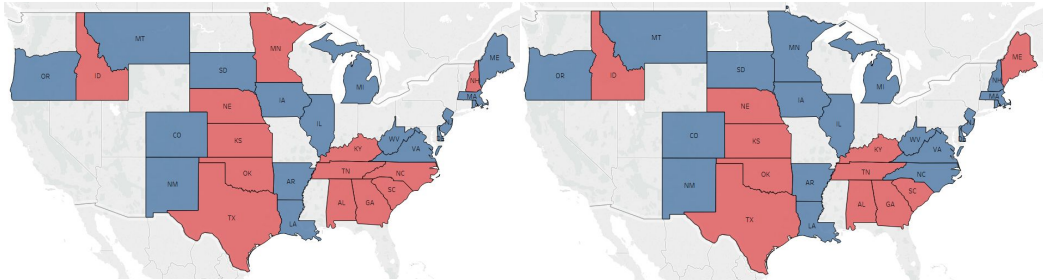
Party Affiliation (2008)



Conclusions

- Identify close elections
- Predict voting outcome solely from donation data





Future Research

Application of other aspects of the dataset

- More recent years
- Other types of elections, including: presidential, House, and primaries
- Effect of independent candidates on the network

Investigate underlying structure

- Co-citation and co-donation
- Preferential attachment

Thank You

- Federal Election Commission (FEC)
- Andrew Waugh
- Professor Porter

Visualizations were made in MATLAB and Tableau