ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

U.S. Senators
A Voting Pattern Study

EE-558 Network Tour of Data Science

Ali Alami-Idrissi, Quentin Bacuet, Leandro Kieliger, Keshav Singh

Demonstration

[Link to Demonstration]

Outline

1. Dataset description
2. Exploratory graph analysis
3. Political affiliation identification
4. Vote outcome prediction
5. Conclusion

Dataset Description

Dataset description

Overview of the U.S. senate

US Senate:

- 100 senators
- Elected for terms of 6 years
- 2 senators per state
- Vote on bills, resolutions and nominations

Decision process:

- Most votes happen "by voice".
- “Roll calls” are recorded , i.e. our data

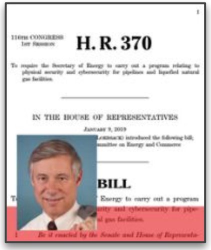
Source: Erin Jackel

Dataset description

ProPublica Congress API

- 129 votes on motions, bills and resolutions.
- 105 senators from 115^{th} congress. Full voting history available for only 96 senators.
- Voting position recorded as numerical value in [-1,0,1]

GovTrack



Bills, resolutions, amendments summaries and latest actions “decrypted”.

Senate and congress websites

Entry point for exploring bills, resolution texts, amendments and roll call interdependence

Exploratory graph analysis

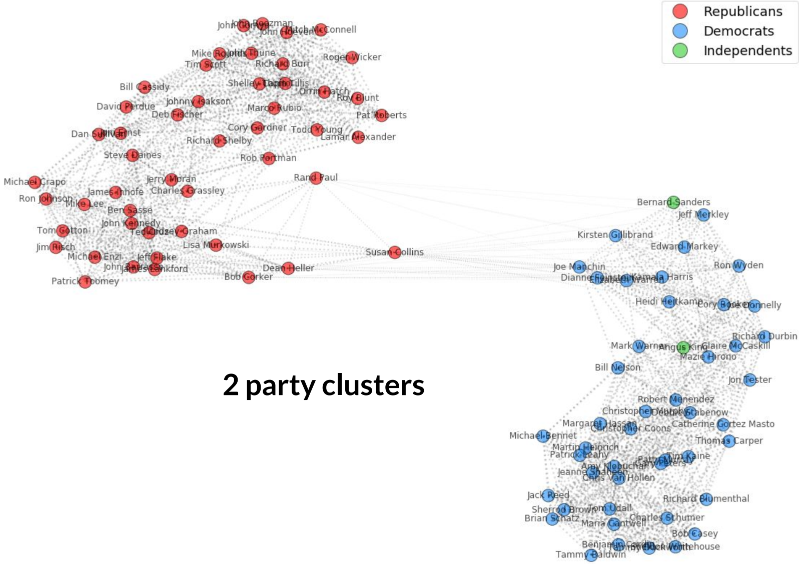
Exploratory graph analysis

Similarity graph construction

- **Nodes:** senators
- **Links:** cosine similarity

$$\delta(\mathbf{x},\mathbf{y}) = \frac{1}{2}\left(1 + \frac{\mathbf{x}^\top \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}\right)$$

- **Sparsification method:**
    - Weight Thresholding at 0.25
    - Limiting the maximum number of neighbors to 20 per node

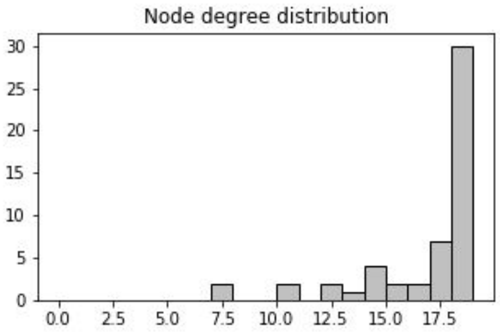


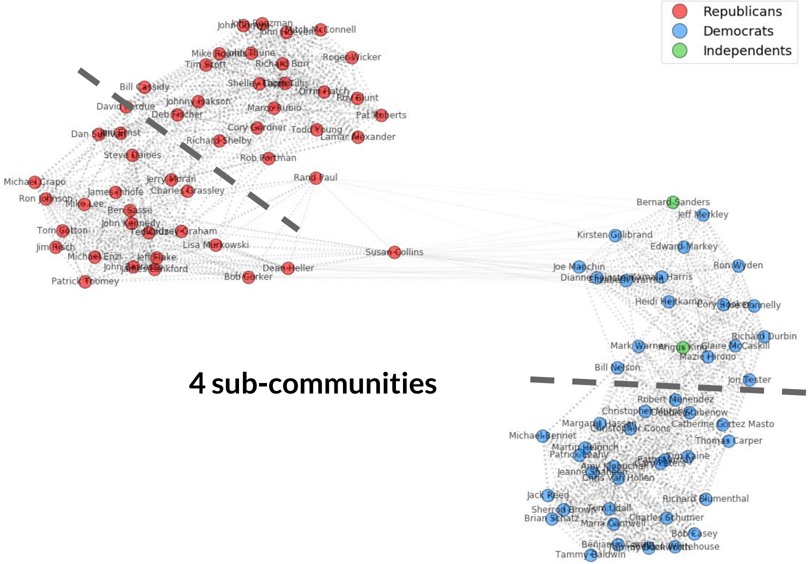
2 party clusters

Exploratory graph analysis

Graph properties

- Clustering coefficient 0.64
- Graph diameter 6
- Single connected component





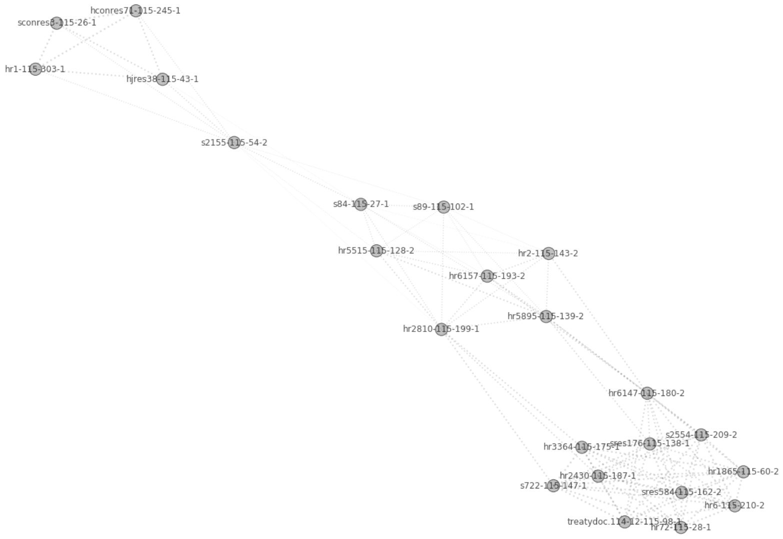
Political affiliation identification

Finding senators that match your ideology

Votes selection

A similarity graph for bills and resolutions

- **Nodes:** Roll-call votes for bills and resolutions only (more interpretable for demo).
- **Links:** Similarity in voting positions computed using cosine similarity
- **Sparsification method:**
  - Weight Thresholding at 0.7
  - Limiting the maximum number of neighbors to 10 per node



Vote selection

Select roll calls which convey the most "political information"

Different techniques used:

- Variance maximization
- Neighbor minimization
- Spectral clustering
- PageRank

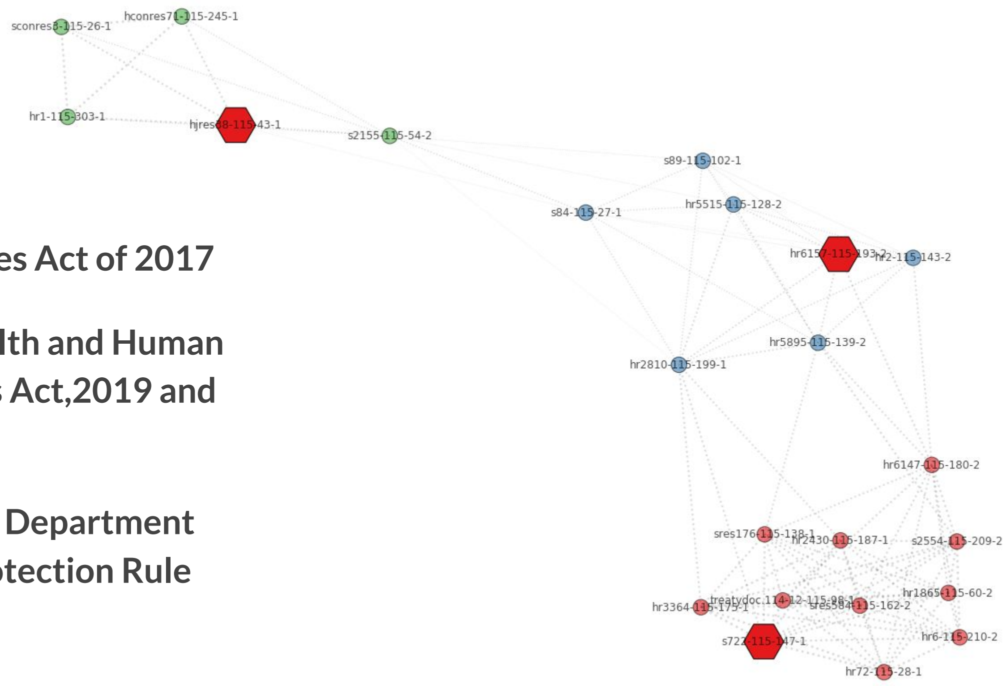
Performance assessment metrics:

- Interpolation accuracy with graph variation minimization
- Interpolation accuracy using K-nearest neighbors

Method	Mean accuracy
Intra-party var. max	0.60
Min neighbors	0.60
Var. maximization	0.61
PageRank	0.88
Cluster centroïds	0.91

Votes selection

Selected votes
(spectral clustering)



1. Countering Iran's Destabilizing Activities Act of 2017
2. Department of Defense and Labor, Health and Human Services, and Education Appropriations Act,2019 and Continuing Appropriations Act, 2019
3. Disapproving the rule submitted by the Department of the Interior known as the Stream Protection Rule

2D Embedding

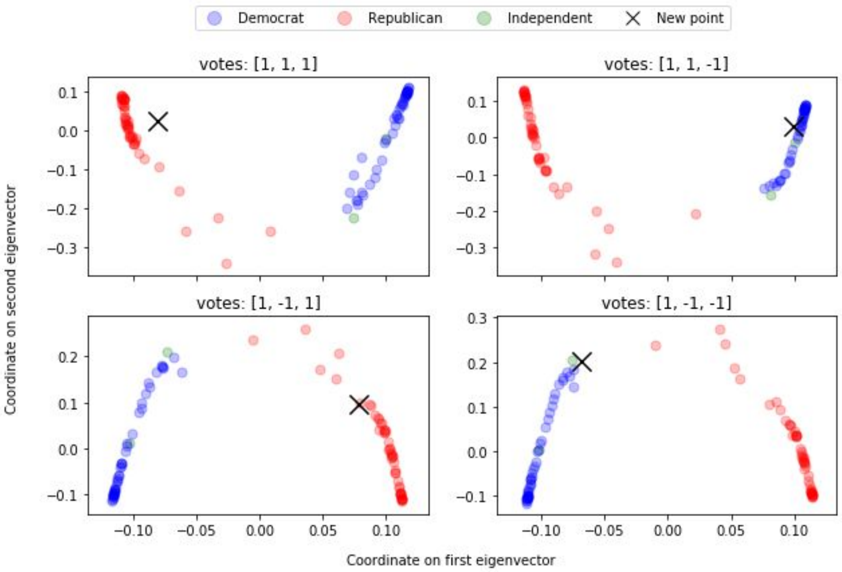
Influence of answers on one's position on the graph

First two votes were passed by a large majority in congress.

➔ Voting against move us away from party clusters

Last vote was well balanced between republican and democrats.

➔ All else being equal, one's position on that issue flips party side



Vote outcome prediction

Using swing senators to interpolate positions across the similarity graph

Vote outcome prediction

Swing senators connect the party clusters

Republicans:

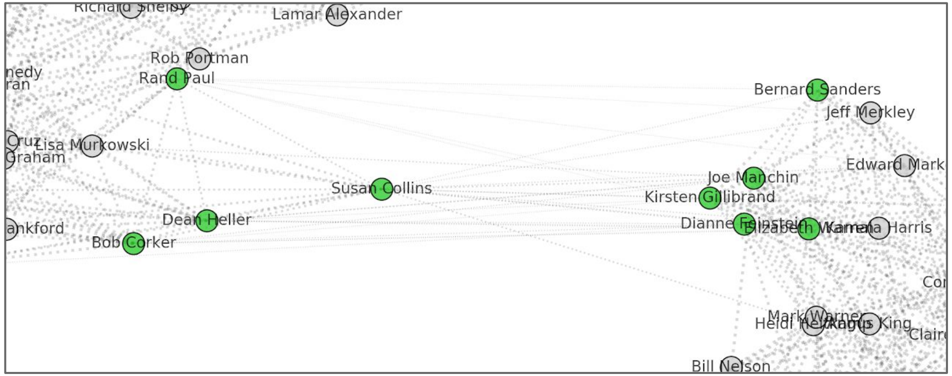
- Susan Collins
- Bob Corker
- Rand Paul
- Dean Heller

Democrats:

- Dianne Feinstein
- Kirsten Gillibrand
- Joe Manchin
- Elizabeth Warren

Independent:

- Bernard Sanders



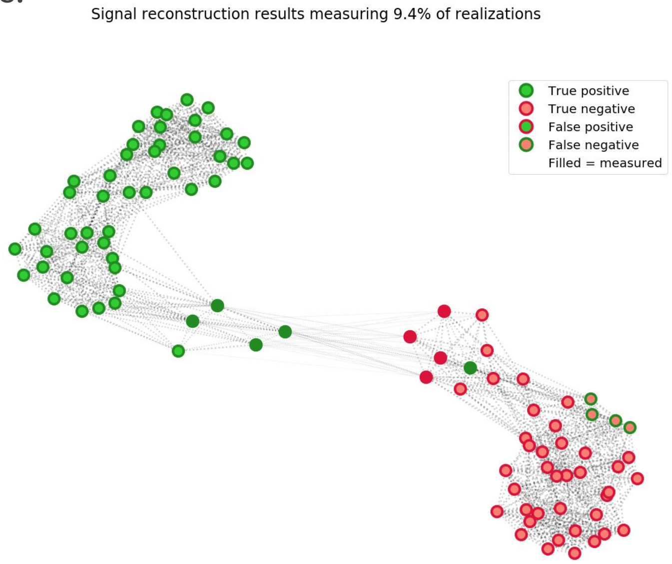
? Can we use them for predicting the outcome of a vote

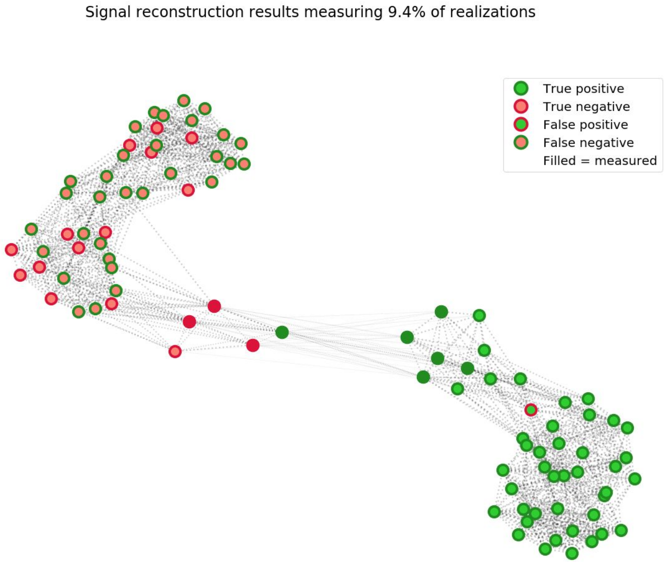
Transductive learning

Interpolation by graph total-variation minimization

$$\min \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sqrt{W_{ij}}|z[i]-z[j]|\right)$$

Solve the minimization problem under the constraint that z retains the voting positions of swing senators.





Transductive learning

Benchmark

Comparison against uniform random sampling (also with 9.5% proportion)

5-fold cross validation. Split set of votes into two sets with proportions 80:20. First set used to build the similarity graph, second one for measuring the accuracy.

Measures	Mean	Std. dev.
swing senators	0.897	0.031
uniform random	0.931	0.023

Conclusion

Political affiliation identification

- Three representatives roll calls help us identify your political stance.
- Spectral clustering approach works best with accuracy ~ 91 %.
- Predict your stand on votes using transductive learning

Vote outcome prediction accuracy

Accuracy of ~ 89 % achieved with Swing Senators and ~ 93% with uniform sampling.

Possible causes of discrepancy:

- Swing senators *can* be more influential with tight votes but there's a lower proportion of tight votes.
- Uniform sampling works well across the board, graph min-cut is already effective at locating majority separation.