UN voting coalitions

Course Project

Shuklov Yuri, Artemii Stetskevich, 2022

The Problem

And implemented idea

- In general, we can model the voting patterns of UN nations in various ways.
- The one chosen for this project is to gather the voting records of those nations into a single complete graph that displays correlation between those records.
- This graph can easily be processed by normal means of community detection and this alone will give us sufficient information on political alignment of the nations.
- The question is: can we process this graph in a manner that enables community detection algorithms to work better.
- The idea attempted in the project: to clean up edges in a manner that ensures proper community split.

The Model.

And it's flaws

- Our problem is a selection of edges to remove from graph $G(V, E)$.
- Normally speaking it is a combinatorial problem, with possible greedy implementation of just tossing out the edges with the smallest weight.
- We propose a slightly more complicated model, based around optimising the following non-linear function:

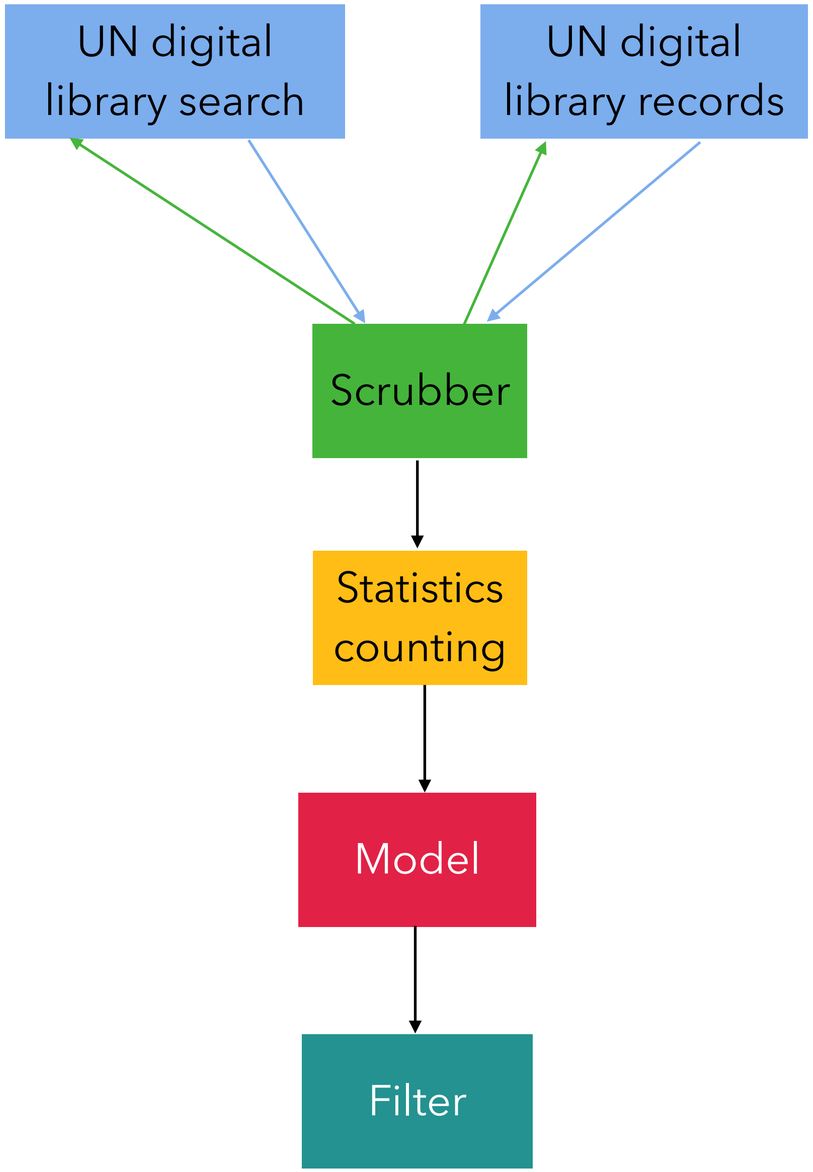
$$\mathcal{F} = \sigma \ln \left[\frac{1}{|V|} \sum_{u \in V} \exp \left(\frac{1}{\sigma} \sum_{v \in V} E(u, v) p_u^T p_v \right) \right]$$

- Where $E(u, v)$ is an appropriate correlation-based weight (the course project uses correlation itself, with correction that $E(u, u) = 0$), and p_u is a distribution of probabilities for vertex u to be in various blocs.
- It's clear that if we find appropriate minimum for this model, we found a partition of nodes into blocs within each of which it is 100% certain that they do not belong in the same community, and as such can toss edges from within without thinking twice.
- The flaw: admittedly, the Lagrange function of this problem is very easy to fall victim to local minimum of uniform distributions. This gives birth to another flaw: the normal descent algorithm must be made deliberately less converging to actually converge on a useful set of distributions.

Implementation

Of things involved

- Scrubber searches for voting records on the UN site, and then goes through each found record for the actual votes cast.
- Gathered records are fed into procedure that computes covariances, and optionally a function of it to produce our graph.
- The produced graph is fed into our model to compute indsets to create. SciPy's interior point algorithm is used for it.
- At last, a separate procedure cleans up the graph from edges in those designated indsets.



Results

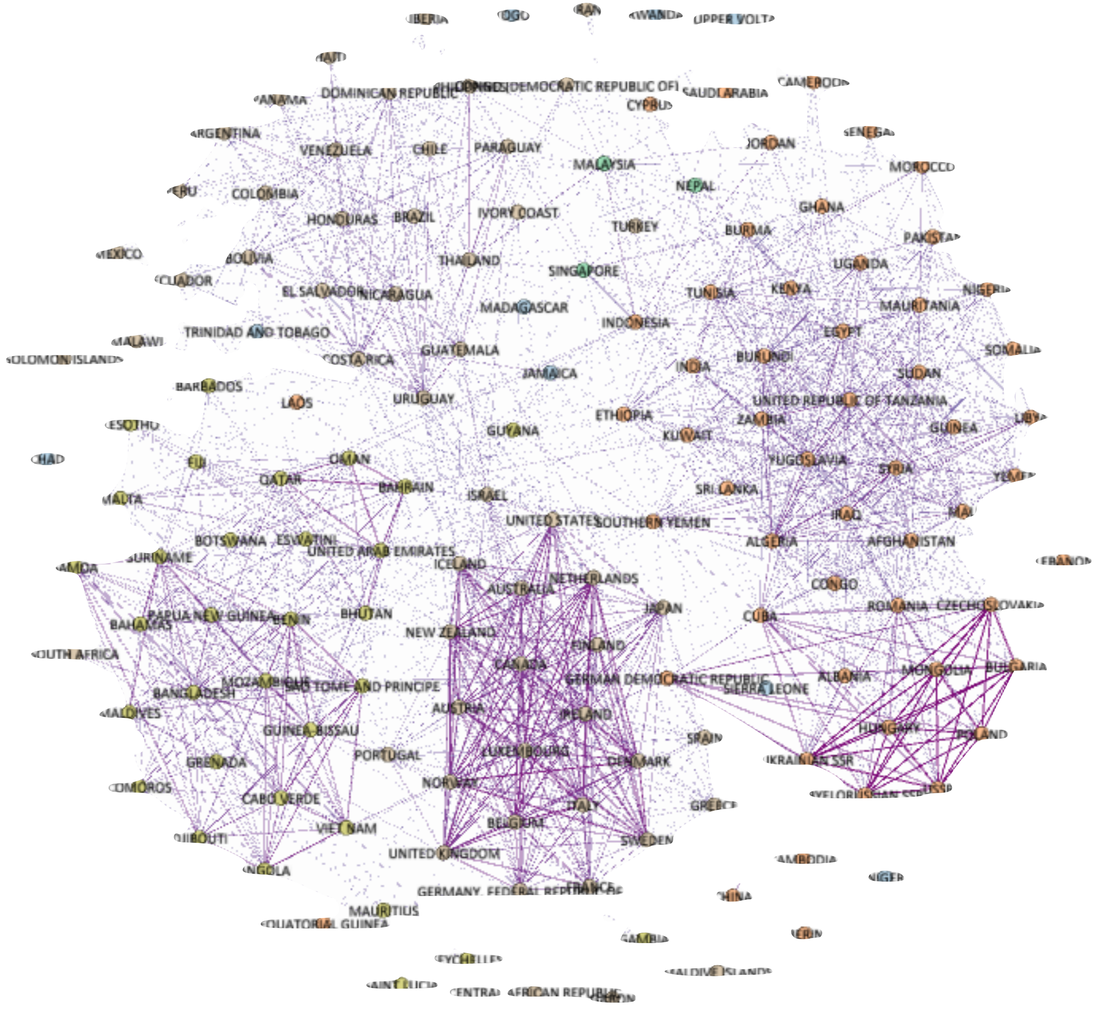
Validation dataset: Cold War records

- A validation to ensure our model does not break things.
- A dataset from years 1965 to 1980, the cold phase of Cold War.
- Has explicit and stable political alliance: Eastern bloc
- If our model does not break it, we may consider it functional.
- As a statement of fact, it does not. The major changes are listed below to communities detected at chosen parameters

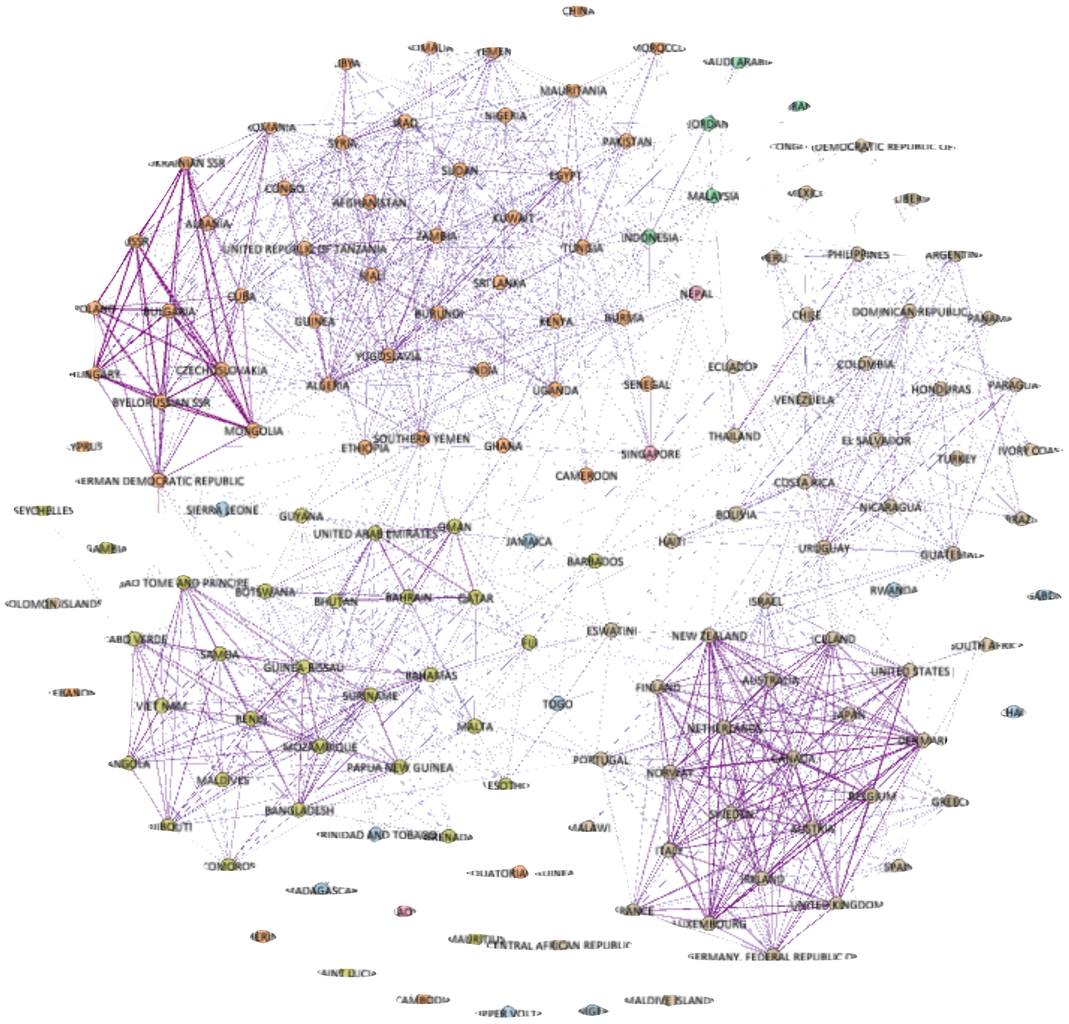
Raw graph	Processed graph
Fiji in community 3, Barbados in community 8	Fiji and Barbados in community 5
Jamaica and Madagascar in community 5	Jamaica and Madagascar in community 4

Results

Validation dataset: Cold War records



Raw graph with minor edges removed



Processed graph with the same filter on edges

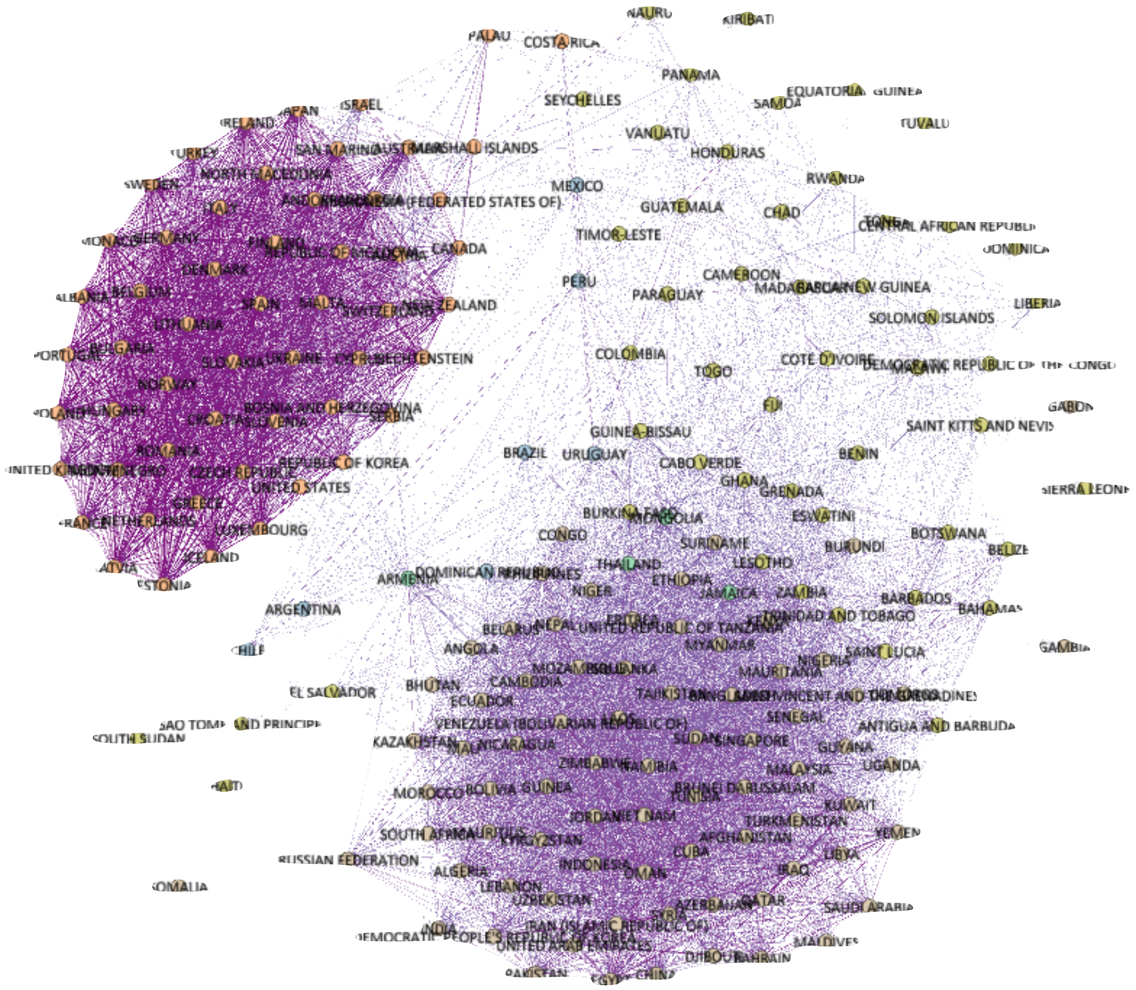
Results

Main dataset: modern day records.

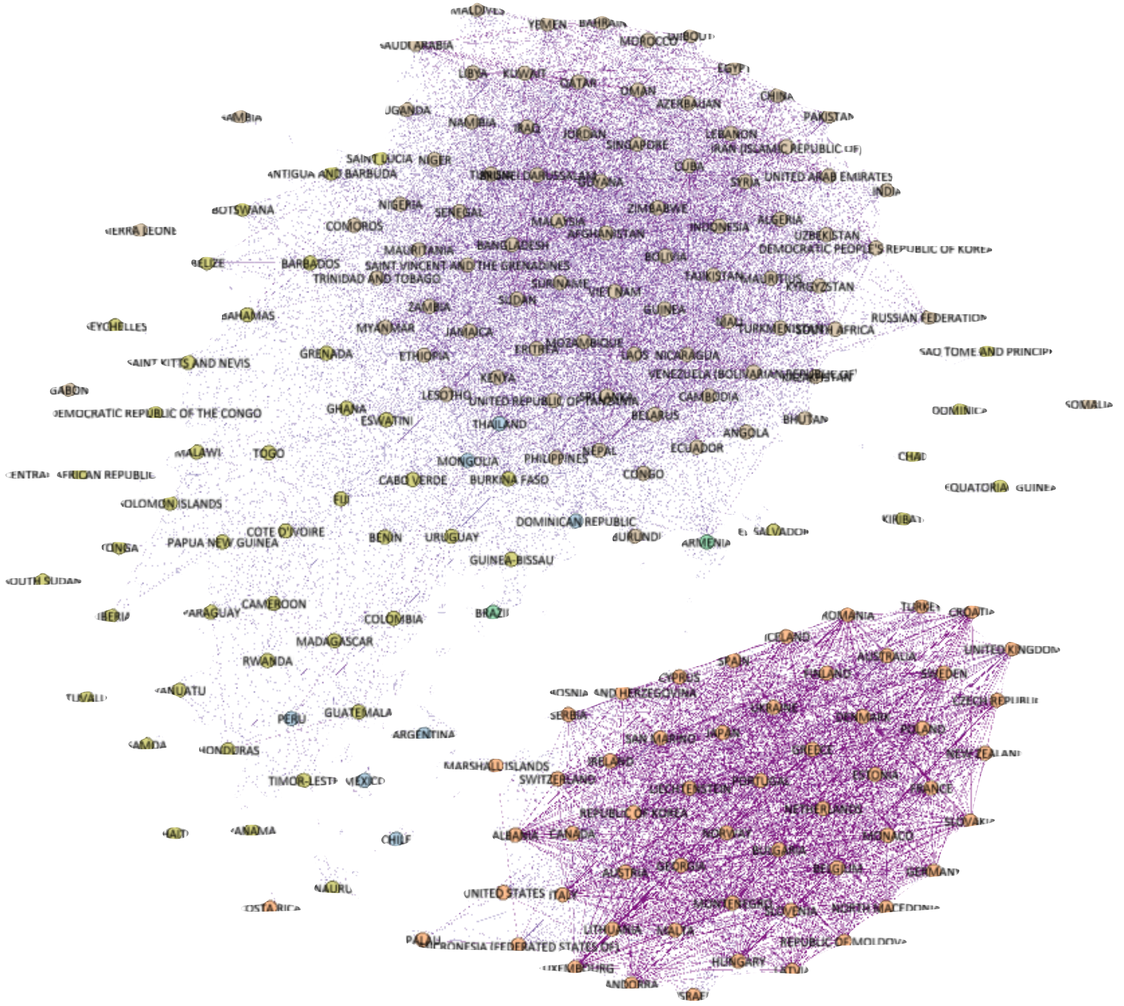
- The main dataset of this problem is a set of votes in modern day, namely from 2007 to 2022.
- We now want to look, whether our model allows to identify things not clear before hand.
- Empirical results suggest yes, albeit with subtle identifications like Armenia-Kahakhstan and Russia-Belarus, not clear at all on original graph

Results

Main dataset:



Raw graph with minor edges removed



Processed graph with the same edge filter

Conclusions

- Model proposed has an effect, but ultimately this effect is too minor to justify the sheer computational cost it has.
- That said, one may make an alternative observation that model proposed is an effective filter on our graph that preserves larger communities while removing the higher weight edges which are arguably more cluttering to the eye.