

Diachronic analysis of co-occurrences networks: a case study on Staribacher diaries and Austrian politics

Matthias Schlögl¹, Marcella Tambuscio¹, Maria Mesner², Matthias Trinkaus³

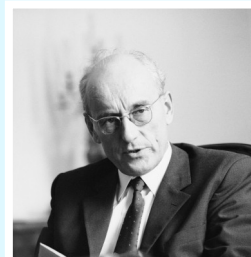
¹Austrian Centre for Digital Humanities and Cultural Heritage, Austrian Academy of Sciences, ² University of Vienna, ³ Kreisky Archives



DATASET



Josef Staribacher (1921-2014) was an Austrian politician, member of the Social Democratic Party and Minister of Trade, Commerce and Industry in 4 governments led by Bruno Kreisky. **He rigorously described in his diaries nearly any meeting and discussion he took part in during his 13 years (1970-1983) as minister:** a very large (about 15000 pages) amount of text that has been recently digitised.



THE NETWORK

We built graph from co-occurrences of people mentioned in the text, using the parse tree of a sentence to find persons who are semantically related.

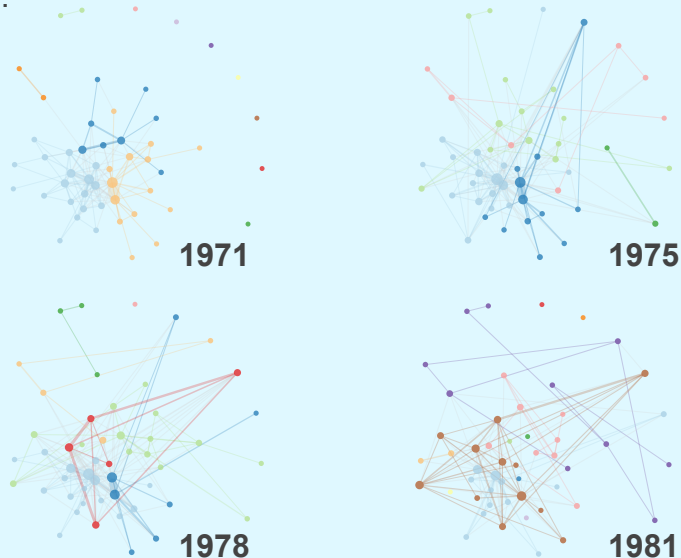
Type of network: undirected, multigraph

Vertices: ~5K, Edges: ~70K

Edges attributes: date (of appearance in the diaries), list of words among the two occurrences

DIACHRONIC ANALYSIS

1. For each year appearing as edge attribute in the network (1971- 1981), we extracted the related **temporal sub-multigraph**.
2. We projected each multigraph to a **weighted simple graph** where edges have an attribute weight that measure how many links connected that couple of nodes.
3. We considered only the nodes that appear in all graphs (years) obtaining a **sequence of 11 graphs** with 51 vertices and a number of links that varies between 82 and 165.
4. The graphs show **appear to be clustered** and exhibit **disassortativity**, then we ran community detection (Louvain): **communities** (more than 6 every year) **change over time**.



TEMPORAL COMMUNITIES

How strong are the connections within the communities during time?

We built another weighted network:

Vertices : 51 people appearing in all the graphs (years).

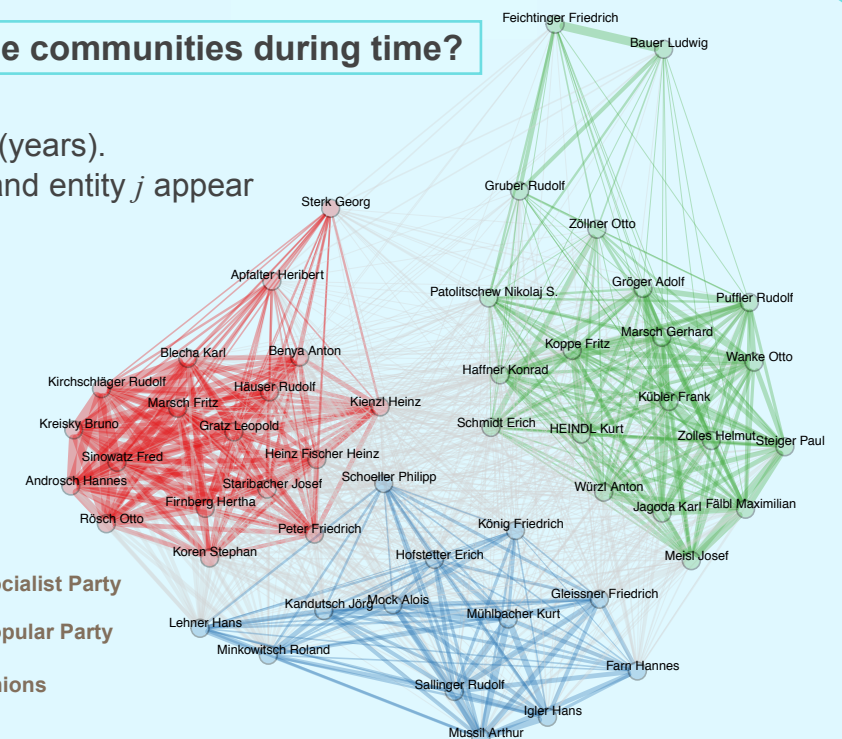
Edges: there is a link (i, j) with a weight w_{ij} if i and entity j appear in the same community in w_{ij} yearly graphs.

Community Detection → 3 communities

The communities reveal quite well the political alignment of involved people: interesting because the list of people mentioned in the diaries does not include only politicians.

Theoretically, this result shows that considering the **temporal evolution of communities** can indeed **improve clustering techniques**.

● Socialist Party
● Popular Party
● Unions



FUTURE WORK

- Adding *institution* label as node attribute: is there an overlapping among the belonging institutions and the detected communities?
- Can the communities be useful to label the nodes that appear less often (not every year)?
- Are the temporal modications in the networks somehow related with some particular event?
- Using the words among the occurrences for further analysis, labeling the relations within topics and exploring their diachronic semantic evolution.

