

# Political Manifesto Analysis

## Technical Report

Giovanni Misseri, Mohammad Kameli, Jolan Chabbey

Network Science course

2019, January

## 1 Introduction

In this relation we will try to explain and justify what we've done in order to analyze the european political society. As one can imagine the political world is dense of alleances, sometimes deleared and sometimes unspoken; and even a bigger problem is the one of understanding the actual political thought of a party. Usually each party declears itself as lined up on right center or left side, but in practice often it's really hard to define what makes a party a right side or left side party and if we just look at how a certain party acts the definition becomes even more fuzzy.

What we aim to achieve with this work is to derive and analize the network of european parties in a cross-temporal analysis. As we said is not always the case that a party that declares itself as a left side party, acts like a left side party. Due to this reason to construct the network in which each party is a node, we will link two parties not if they declare to collaborate each other and nor if they have some political alleance, we will try to link two nodes if the two connected parties have the same idea on some critical concepts.

On such network we will run our analysis trying to spot behaviours and

insights of the political scene. Moreover we will try to detect the communities inside this network in order to see if the actual division of the parties overlaps with the declared one, and consequently see which are the parties that, based on their ideas, are not distinguishable by their adversary.

## Manifesto Project

The Manifesto Project - "<https://manifesto-project.wzb.eu/>" - is a German project which aims to maintain and continuously update different political datasets. Their main work is the "Manifesto Project Dataset", a dataset containing at the moment 4388 political manifestos published from 1945 till today by 1117 parties. They publish and maintain this dataset in order to be able to conduct scientific research on the European political field, but not explicitly exploiting network science or graph theory.

We found this project really interesting because it contains information about the processing of parties' political manifestos. The great advantage of using the Manifesto Project work is not only the high number of manifestos present there, but also the way they extracted information about manifestos and how they encoded that. In the Manifesto Project dataset there is a good number of variables representing different ideas, like "Imperialism" or "Welfare State Expansion", giving them for each manifesto a score representing "how much a manifesto reflects that idea".

As we said we wanted to build a network of parties exploiting what they think and not what they declare, here the limit is evident but building a network on political manifestos and the ideas that it reflects could be considered a good approximation of reality.

We will then exploit the Manifesto Project dataset to build our network.

## 2 The Network

The first problem we had to deal with is how to exploit a general dataset to build a meaningful network. We come up with a nice idea and after some research we found out that it's a quite used way to proceed. Before talking about how we built our adjacency matrix it's important to underline the data engineering we went through.

Our dataset was composed by 1117 parties and we had in total 4388 manifestos, so we had an average of 4 manifestos for each party. We had to decide how to deal with the greater number of manifestos with respect to the parties. A natural way to proceed is to average the score of the manifestos for each party, and that's what we did. This is a meaningful way to proceed because like that what we work on is an average idea's score over the time for each party. Another possible way to proceed would be taking only the last available manifesto's score, one would have obtained the most recent data but losing past information, this could bring to biased measures due to the variability of ideas in parties, a mean over time brings to zero the score for all that ideas where a party showed contradictory positions over time, underlining the one where a party actually based his existence.

Said that we can jump to the adjacency matrix. We took our restructured dataset made by 1114 rows, the number of parties without missing data in target variables, and 56 columns, the set of main ideas Manifesto Project gives score. First of all we computed the pairwise distance matrix, in which in position  $i,j$  is present the euclidean distance between party's  $i$  ideas and party's  $j$  ideas. At this point we needed to decide a threshold for which if two parties are nearer than that, they are connected, otherwise they are not. This approach will create a symmetric unweighted adjacency matrix and will be the one on which we will conduct the greatest part of our analysis. A particularity of this approach is that is comparable with the hierarchical clustering in particular with the single linkage hierarchical clustering, this will be crucial for future analysis.