

Lockdown Mobility Network Analysis

Final project for the Social Network Analysis course

Gabriele Sarti

Enrico Fallacara

University of Trieste, A.A. 2019/2020

Abstract

The aim of this project report is to analyze shifts in mobility patterns during Italy's COVID-19 lockdown. More specifically, we consider the Italian mobility network for three specific days during the COVID-19 pandemic outbreak: the beginning of the **first local lockdown** in the Lodi province (25/02/2020), the beginning of **nation-scale lockdown** (03/10/2020) and the **end of nation-scale lockdown** at regional level (05/05/2020). We perform a diachronic descriptive analysis of the mobility network, apply community detection algorithms to highlight mobility shifts at a provincial level and model network connections by the mean of exponential random graphs (ERGs) to show the strong connection between mobility shifts and COVID-19 diffusion statistics. We conclude by comparing our results with the mobility reports publicly released by Google during the lockdown. Our code is available at <https://github.com/gsarti/lockdown-mobility-analysis>

Introduction

The nation-level lockdown implemented by the Italian Government in order to prevent the spreading of the COVID-19 disease in 2020 greatly affected the mobility patterns of roughly 60 millions of Italian citizens and residents. The Italian lockdown was composed by several subsequent phases of tightening and loosening of restrictions, with three most relevant dates arguably being:

- The **first lockdown** of ten municipalities of the province of Lodi, forming the first *zona rossa*, shortly after February 21st, 2020.
- The **national lockdown** that restricted travel at intra-municipal level, except for working necessities and family emergencies, on the evening of March 9th, 2020.
- The **reopening of inter-regional travel** that took place from May 5th, 2020.

In this report, we analyze how the Italian mobility trends have changed during the three phases defined by the aforementioned dates using a dataset containing raw movements *WAIT FOR SPECIFICATIONS* at a macro-scale level. We begin by presenting data and preprocessing steps that were taken in order to perform the analysis in [Section 1](#). [Section 2](#) contains a diachronic descriptive analysis of the mobility network that highlights shared and different properties of the networks through time. Inside [Section 3](#), we apply highlight mobility shifts at a provincial level by the means of well-established community detection algorithms. [Section 4](#) contains our efforts in modeling the mobility network by the means of exponential random graphs (ERGs) to show the strong connection between mobility shifts and COVID-19 spreading statistics. We conclude in [Section 5](#) by comparing our results with the mobility reports publicly released by Google during the lockdown and summarizing our findings.

1. Data and Preprocessing

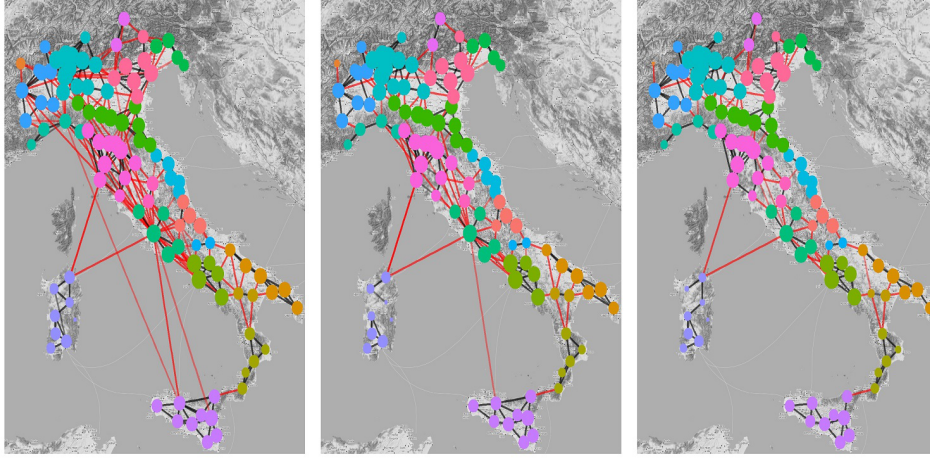
The datasets contains the estimated raw number of movements n in 3 days during the Covid crisis, filtered by country = IT. The 3 days refers to In this dataset are present different metric to estimate movements, we have selected the metric n because it represents a kind of measure of popularity.

2. Diachronic Descriptive Analysis of Mobility Networks

A first insight is given by the sum of the movements in these 3 days. We notice that mobility doesn't change much allowing intra-provincial movements, since inter-province connections have values that are much lower than intra-province ones. Thus, removing the loops, allow us to see only inter-province movements and in this case the mobility is significantly lower:

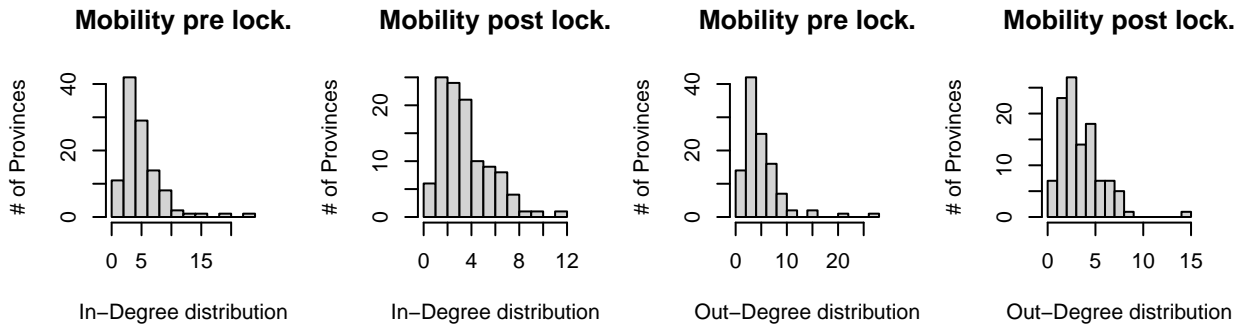
```
## [1] "Movements (inter-province): Pre 119065 -> Mid 85258 -> Post 68515"
```

A graphical representation of this phenomenon is given by the next figure.



We restrict our attention to inter-province movements for the descriptive analysis, since they capture better the variation of mobility and the results for intra-province are empirically quite similar. The graphs are **directed**, **weighted** (according to the metric value) and **disconnected** (this is valid for both implementations).

From the in-degree and out-degree distribution of the 3 graphs, it is highlighted that the majority of nodes have degree between 1 and 10, with few nodes that assumes bigger values. We can notice that, for in-degree, large values tends to vanish and the distribution concentrates to small values, indicating that nodes loose their popularity through time. The same happens to the out-degree distribution, indicating that also expansiveness of nodes tends to reduce.



The **network density**, presented in Table 1, indicates, first, that the graph are **sparse**, since the very low values that assumes this quantity, and second, that the denisty reduces over time, indicating, like the in-degree and out-degree distributions, that the movements significantly decreases between the provinces through time.

In order to implement centrality measures, we have to invert the weights of our graphs. This have to be done because actual weights represent a sort of strength measure between nodes, not a distance(cost) measure. After that we can calculate the **betweenness** and the **eigenvector centrality** for the graphs. Note that the closeness will be calculated later, because this measure is only meaningful in a connected graph.

Table 1: Descriptive analysis measures

	Pre-lock.	Mid-lock.	Post-lock.
Density	0.0500417	0.0398666	0.0361968
Assortativity by degree	0.1343682	0.2511714	0.3258981
Transitivity Glo.	0.3863636	0.4052718	0.4171598
Transitivity sub. upper	0.1714286	0.2068966	0.2727273
Transitivity sub. lower	0.3148734	0.3176471	0.3448702
Transitivity k-cores	0.7959184	0.6127321	0.6113208

The betweenness remains practically the same in all the three graphs, apart from a few small variations. However in Lazio, Umbria and Abruzzo happens an interesting thing. Before lockdown, Rome has the largest value with respect to neighboring provinces, but as the time passes, its value tends to reduce and the other neighboring provinces tend to assume higher values. This could indicate that Rome’s tendency to influence movements (flows) is slightly lost in favor of its neighbors.

CITARE PLOT BTWN SUBGRAPH IN LAZIO, UMBRIA E ABRUZZO

Eigenvector centrality shows that there is a loss of influence over the whole network by provinces in the province of northern and southern Italy, while in central Italy the situation remains almost completely unchanged, despite a slight decrease in values.

CITARE PLOT EIGE_VECT SUBGRAPH IN TUTTA ITALIA

For what it concerns about **assortativity by degree**, we calculate this measure for the graphs and, as reported in Table 1, we can see that in the mobility pre-lockdown we have a small positive value of 0.134 that indicates the slight tendency of nodes of similar degree to connect with each others. An interesting fact is that this measure increases through time, reaching a value of 0.326. This could indicate that the movements have concentrated more between provinces of the same degree, reducing the movement from “small” to “large” ones.

The **global clustering coefficient**, that refers to network cohesion analysis, is calculated:

- **Globally**
- on **subgraphs** in which all the nodes have a weight that is above or below a fixed **threshold**.
- in **k-cores components** of original graphs.

As reported in Table 1, the **transitivity** for the whole network tends to remain the same, except for a small increase that isn’t significant. Thus, to inspect better this property, we calculate this measure on two subgraphs obtained placing a $threshold_w = 500$ on weights. The value that we choose is based on the distribution of weights in the network, noticing that most part of nodes have weights that go from 0 to 500. As we can notice from the results presented in the table, for the subgraphs of nodes below the threshold, the transitivity, that is smaller with respect the global value, remains again practically the same, instead for the other subgraph there is a significant increment of 0.10. This could mean that big provinces tends to increase their connections through the three different phases of lockdown. We have also calculated this index on k-cores of the graph, selecting as k the minimum value between the three coreness value, resulting in $k = 8$ (notice that only the graph referred to the mobility pre-lockdown has a value of k bigger than the selected, with the coreness equal to 10). The transitivity calculated on these maximal subgraphs has a significant decrease of 0.18, as we expected, indicating that these subgraphs have a loss in terms of cohesion through the time.

After this first descriptive analysis we can deduce that the decrease of movements caused by lock-down measures causes a loss of influence of the nodes in the whole network, with an increasing tendency of similar degree nodes to connect with each others, indicating that movements from small provinces to large ones decrease, in favor of smaller and closer provinces, in terms of distance. The opposite happens for big provinces, movements from bigger to smaller reduce in favor of movements between big ones.

3. Detecting Mobility Communities
4. Modeling Mobility Patterns
5. Discussion and Conclusion

A. Coreness Plots for Feb. 25th, 2020 Mobility Network

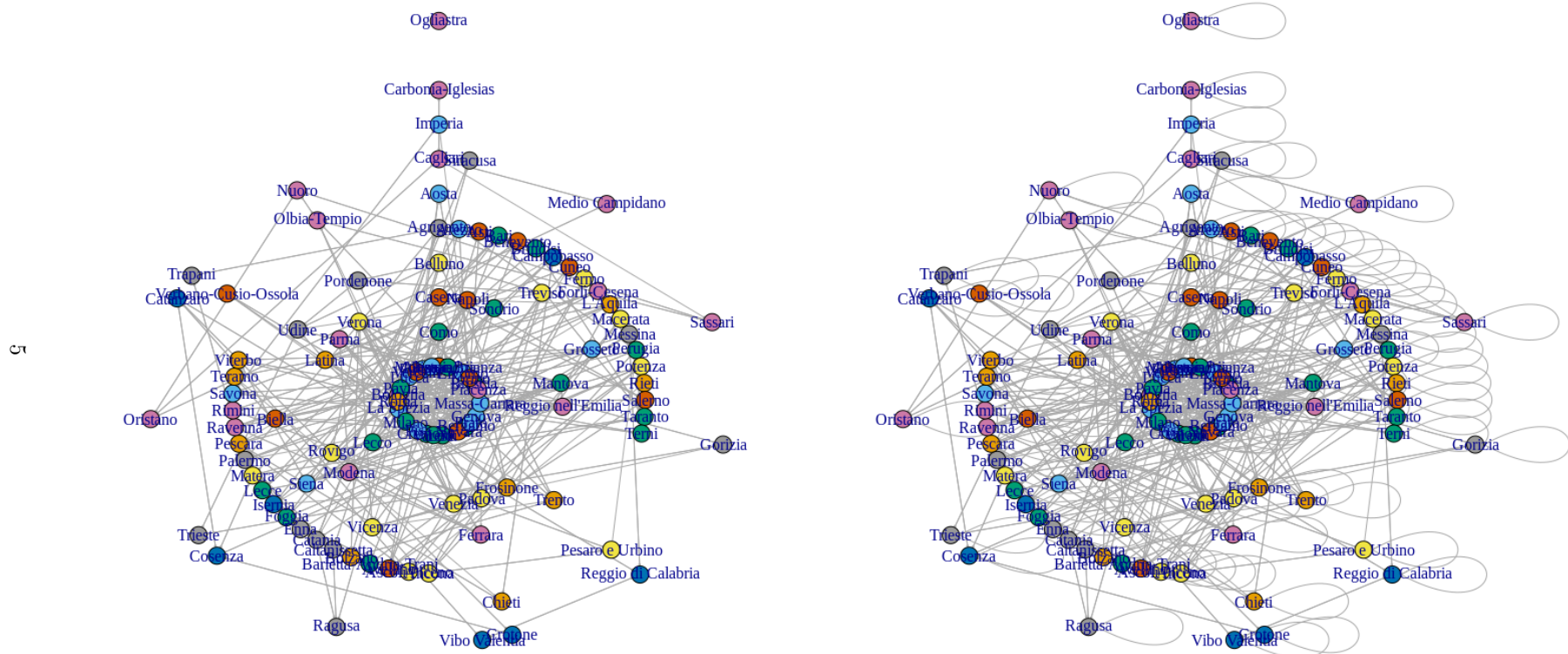


Figure 1: Finding the .rproj file for Chapter 3

B. Weighted Coreness Plots for Feb. 25th, 2020 Mobility Network

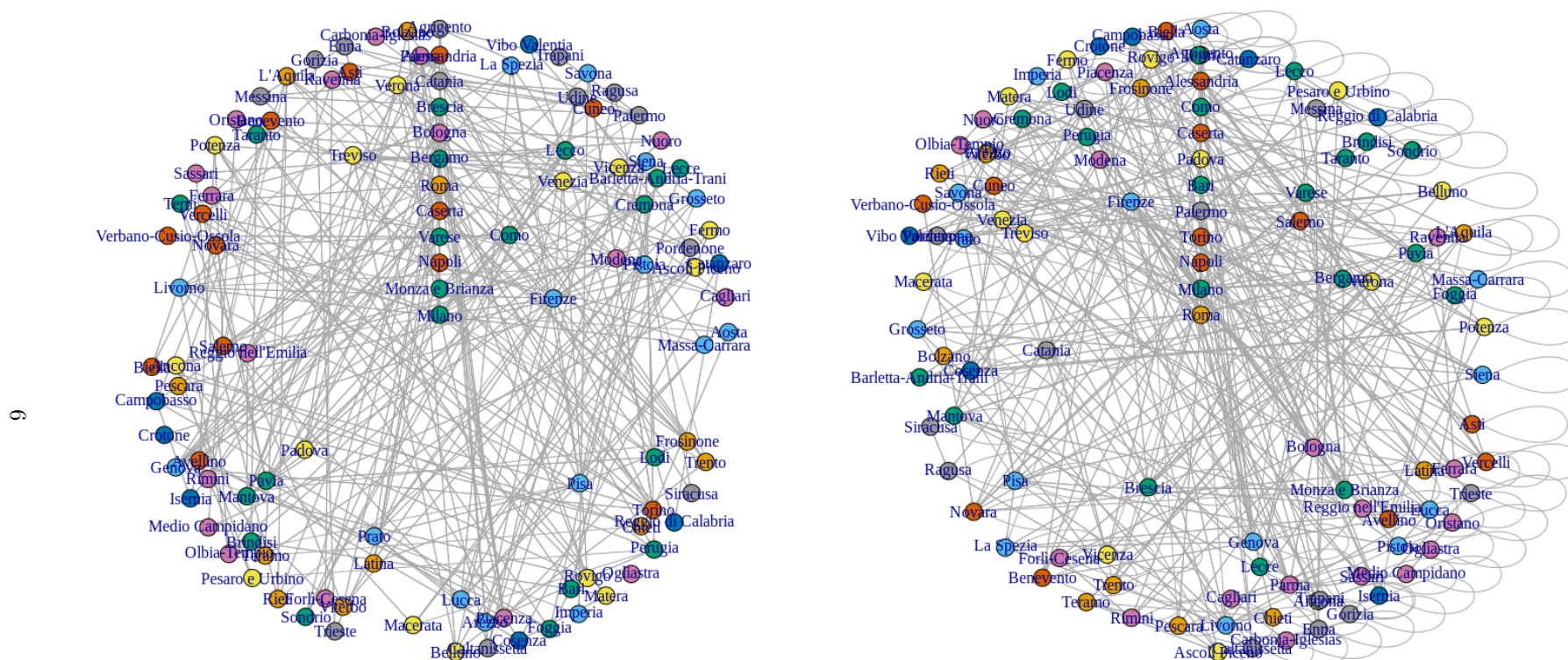


Figure 2: Finding the .rproj file for Chapter 3

C. Connected Components and Cliques for the Three Lockdown Phases

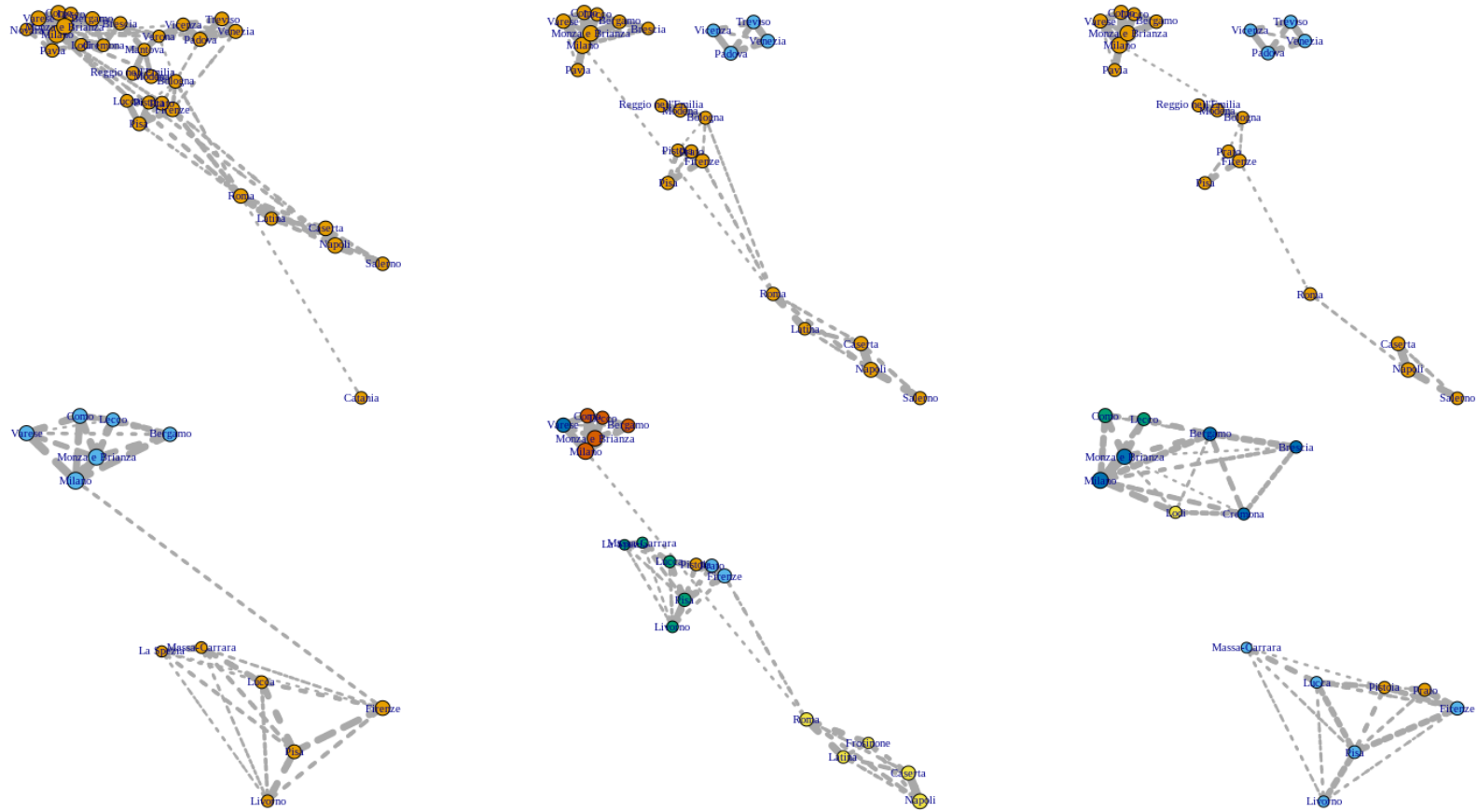


Figure 3: Finding the .rproj file for Chapter 3

C. Community Detection using Girvan-Newman Algorithm and Label Propagation

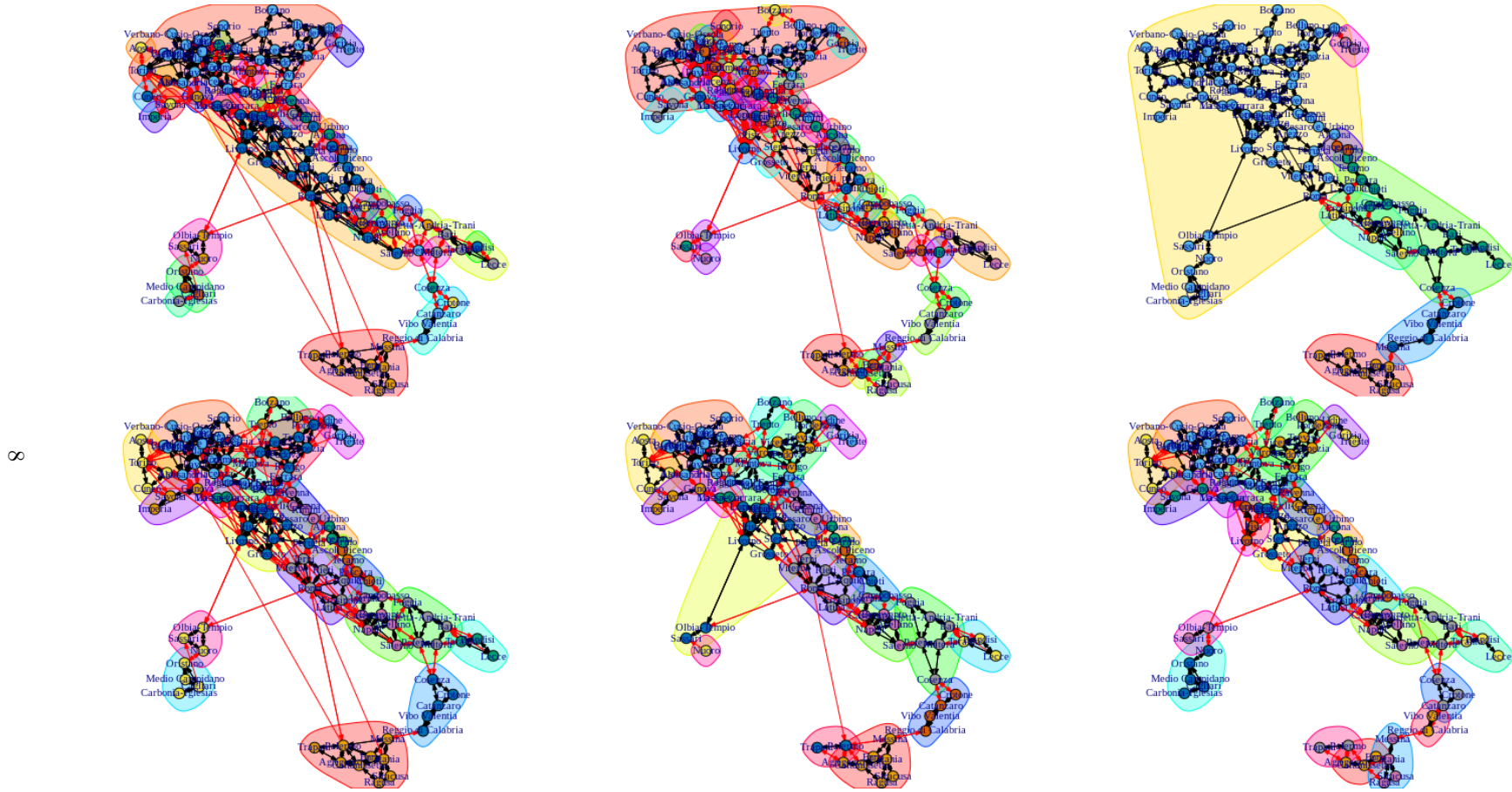


Figure 4: Finding the .rproj file for Chapter 3

D. Regions of middle Italy where the size of each vertex is proportional to betweenness

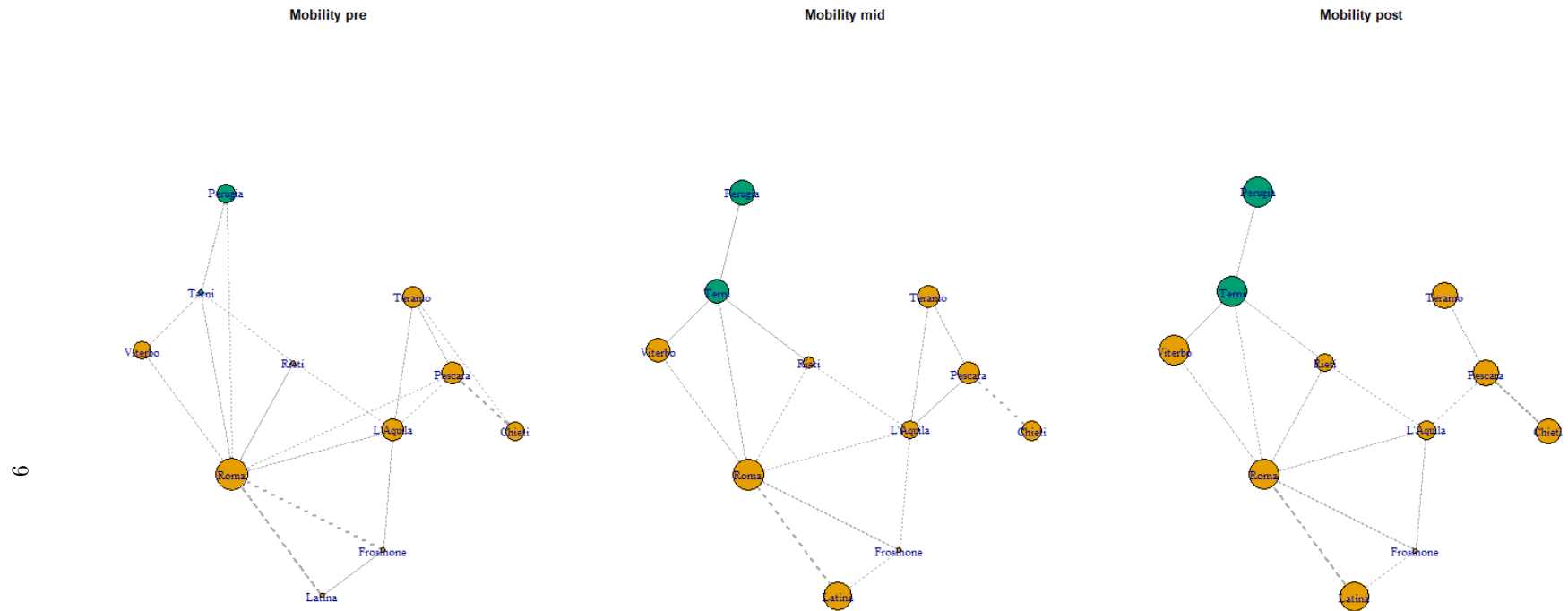


Figure 5: Finding the .rproj file for Chapter 3

10

