Networks – Problem Set 1

Konstantin Boss

January 17, 2020

Exercise 2: The Medici Family Network

In this exercise we are going to examine a network of 15th century Florentine families to determine whether the fame that the Medici family has earned could be explained using network techniques. The reference for this exercise is *Padgett & Ansell* (1993).

Comparing Centrality scores

In this exercise we construct a table that compares three different measures of node centrality for the 16 families in our Florentine network. These metrics are:

1. Degree centrality - how many direct connections does the family have?
2. Harmonic centrality - inverse of what's the average length of the shortest path between the node and all other nodes?
3. Betweenness centrality - how many connections pass through the node?

The results are summarized in Table **??** and show that the Medici family comes out on top in all three categories, followed by the Guadagni family. In that sense they appear to confirm the claim by *Padgett & Ansell* (1993) that the Medici family was central to Florentine society in the 15th century. Moreover, we can see that the second most important family, the Guadagni family, was not as far behind as one might think in terms of degrees or of harmonic centrality. However, the betweenness measure for the Medici family is about double that of the second family. This suggests that the Medici were an important inter-connector within our Florentine network.

Betweenness	Names	Degree	Names	Harmonic	Names
47.50	Medici	6.00	Medici	9.50	Medici
23.17	Guadagni	4.00	Guadagni	8.08	Guadagni
19.33	Albizzi	4.00	Strozzi	8.00	Ridolfi
13.00	Salviati	3.00	Albizzi	7.83	Albizzi
10.33	Ridolfi	3.00	Bischeri	7.83	Strozzi
9.50	Bischeri	3.00	Castellan	7.83	Tornabuon
9.33	Strozzi	3.00	Peruzzi	7.20	Bischeri
8.50	Barbadori	3.00	Ridolfi	7.08	Barbadori
8.33	Tornabuon	3.00	Tornabuon	6.92	Castellan
5.00	Castellan	2.00	Barbadori	6.78	Peruzzi
2.00	Peruzzi	2.00	Salviati	6.58	Salviati
0.00	Acciaiuol	1.00	Acciaiuol	5.92	Acciaiuol
0.00	Ginori	1.00	Ginori	5.37	Lambertes
0.00	Lambertes	1.00	Lambertes	5.33	Ginori
0.00	Pazzi	1.00	Pazzi	4.77	Pazzi
0.00	Pucci	0.00	Pucci	0.00	Pucci

Table 1: Centrality statistics for Florentine network

Configuration model approach

To test whether the importance of a node, in this case the Medici family, can be inferred purely from its degree, we run a simulation of 1000 graphs using the configuration model with the same degree sequence as in the original data and have the available connections rewired at random. As we have an undirected graph, the total of nodes has to be even, which we check and confirm. Then, we determine the harmonic centrality measure which gives us information about the average length of the shortest paths between a node and all other nodes (inverted). Finally, we compare the obtained mean to the measures we found in the exercise before by computing the difference. The result is presented in Figure **??**.

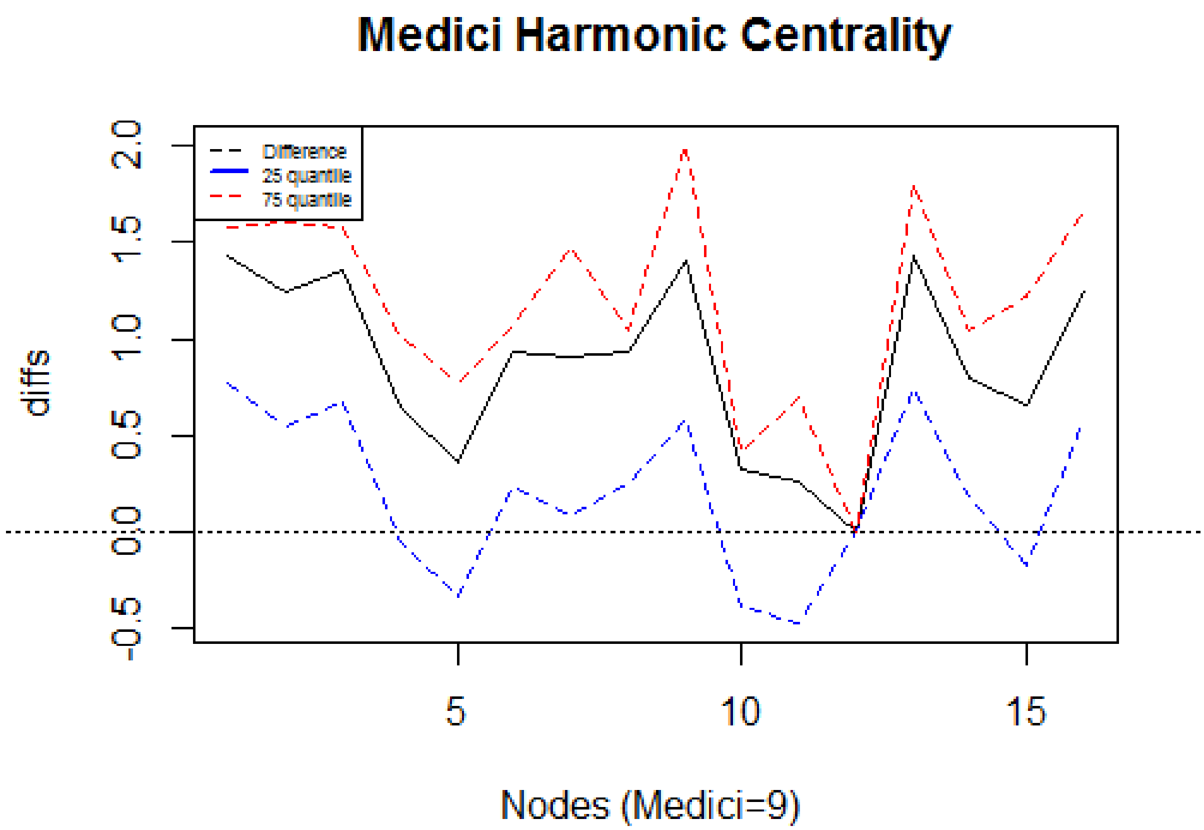

Figure 1: Harmonic centrality metric and quantiles for 1000 configuration models

The figure shows that the difference between the original measure and the mean from the 1000 simulations is always positive, so that the original data possibly represents a favorable draw of the distribution for the Medici family. Moreover, especially at the node of the Medici family the quantiles are quite wide. So we conclude that inferring popularity solely from the degree of a node, respectively from the associated degree distribution may be misleading.

Exercise 3: Villages in India

In this exercise we examine the networks of nine villages in India based on (1) visits between members of each village and (2) on money transfers between inhabitants of each village. For this we first compute the intersections of the adjacency matrices as suggested.

Plotting the network graphs

We finally decided for the suggested Fruchterman-Reingold layout for our graphs. This gives them a good overview between larger and smaller subnetworks. Results are depicted in Figures **??** and **??**.

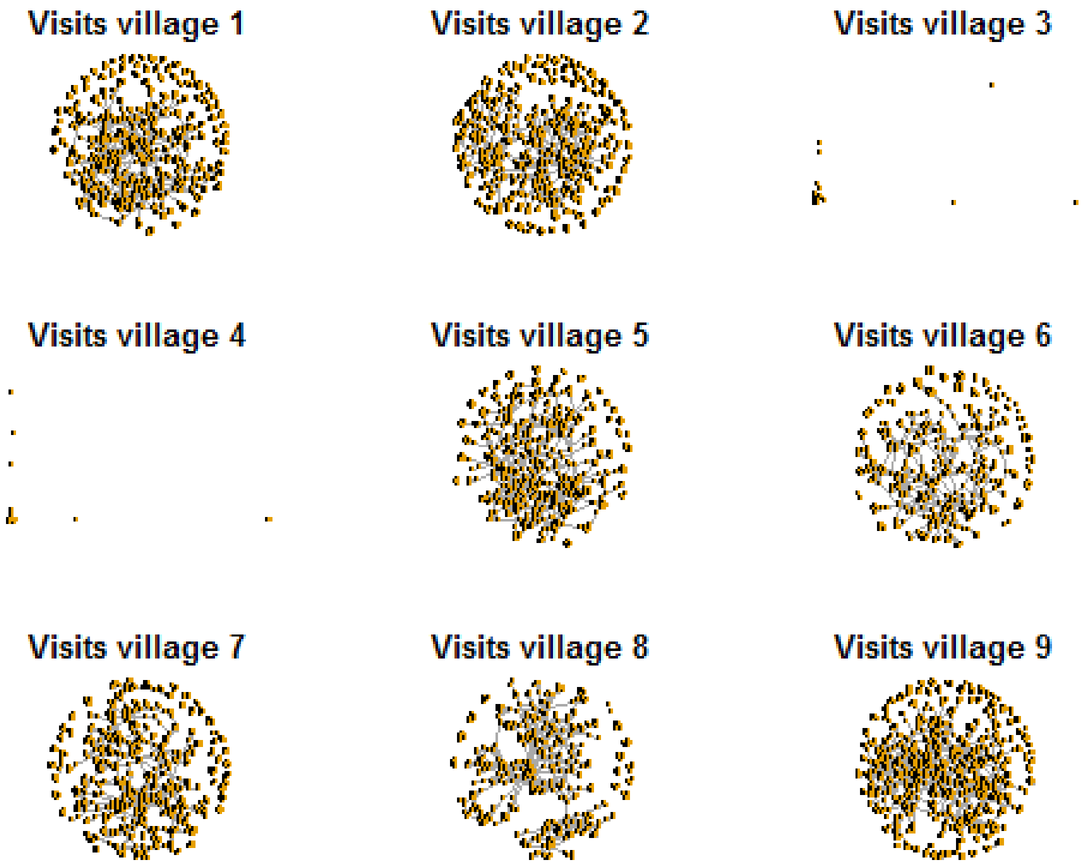

Figure 2: Networks of visits between inhabitants of 9 Indian villages

Figure 3: Networks of money transfers between inhabitants of 9 Indian villages

In both cases (visits, money) we observe that in the 3rd and 4th village there are not many interactions ongoing. This may be a result of different factors: first of all, we may have miscalculated some things and gotten intersection matrices with too many zeros. Second, there may just not be a lot of interaction between inhabitants of these villages which is surprising as these are by some margin the largest villages. Third, it could be a scale problem of the graph, as the distance between individuals could be too large to plot well in this format. Fourth, it could be that interactions are directed and thus would not appear as a 1 in our merged adjacency matrices. Overall, the graphs show nicely that there are usually some smaller factions (depicted at the periphery) in a village and a large connected nucleus.

Descriptive statistics for the villages

We compute six metrics of network relations for the nine villages in both categories which gives a total of 18 lists of descriptive data. The 7th metric, the Bonacich measure, is omitted as the adjacency matrices (except for one) turn out to be not-invertible by the algorithm that is currently used by *igraph*. Due to the sheer amount of data, we refrain from presenting it here. Upon running the code the data can be obtained in the lists labeled *Descr_stat_money_** and *Descr_stat_visits_** respectively, where the asterisk stands for the village number.