Università Ca'Foscari Venezia

NETWORKS IN ECONOMICS AND SOCIAL SCIENCE

Instructor: Prof. Casarin

Student: Hoan Dao (871510) + Duc Duong (871492)

Venice, Italy - 29th April 2018

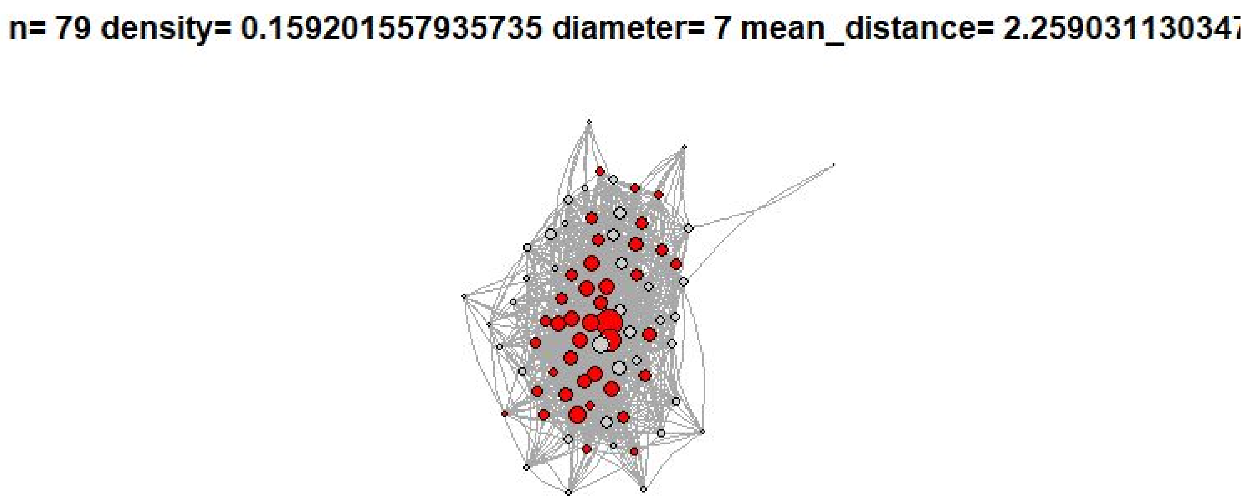
1. Introduction

In this era, a huge amount of money is transferred between banks each day. Understanding about this data can bring us a picture about the relations of these banks and the health of each bank. A good way to do this is applying for graph theory, which can show us how banks are connected.

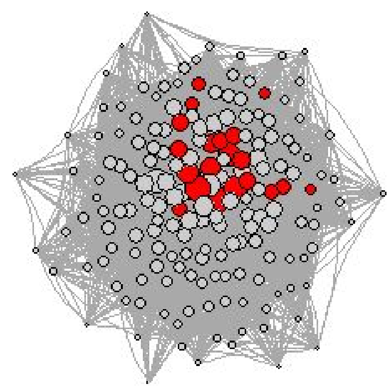
The data set for this project is provided by Prof. Casarin (r.casarin@unive.it) of Ca' Foscari university of Venice. The data set contains information about 210 bank networks which is recorded in different time. Based on this data set, we will calculate some statistic indexes for all networks in the data set to show the evolution over time. We also focus on analysis some special networks to show how banks are connected with each other.

2. Overview of the relations between banks

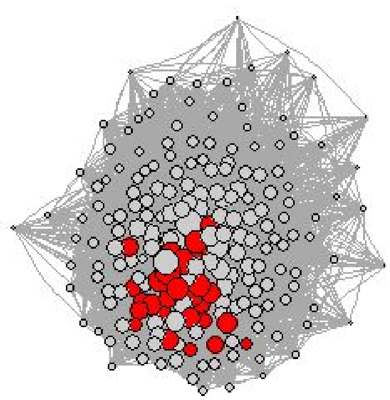
Let pick up the graphs number 10, 40, 110 and 150 in the data set.

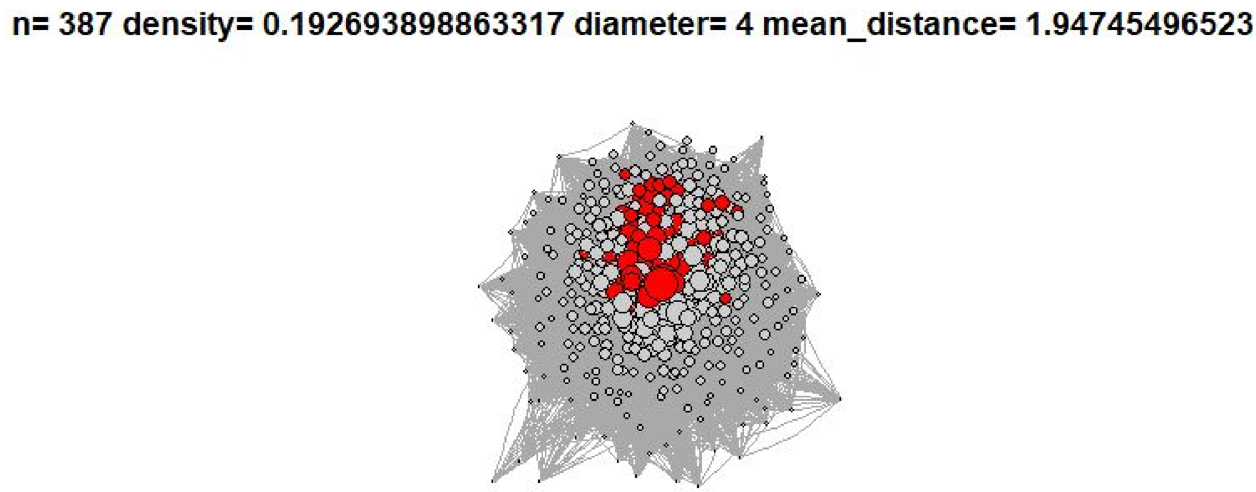


n= 180 density= 0.185567970204842 diameter= 4 mean_distance= 1.92774674115



n= 201 density= 0.196517412935323 diameter= 4 mean_distance= 1.86131840796

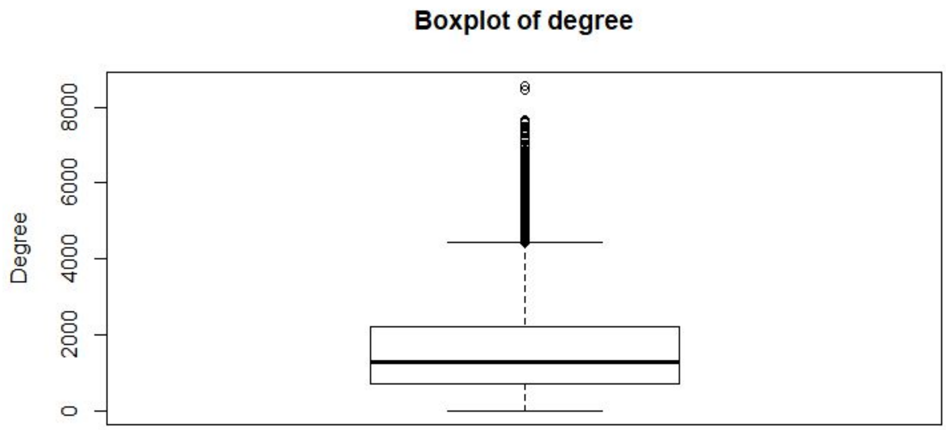




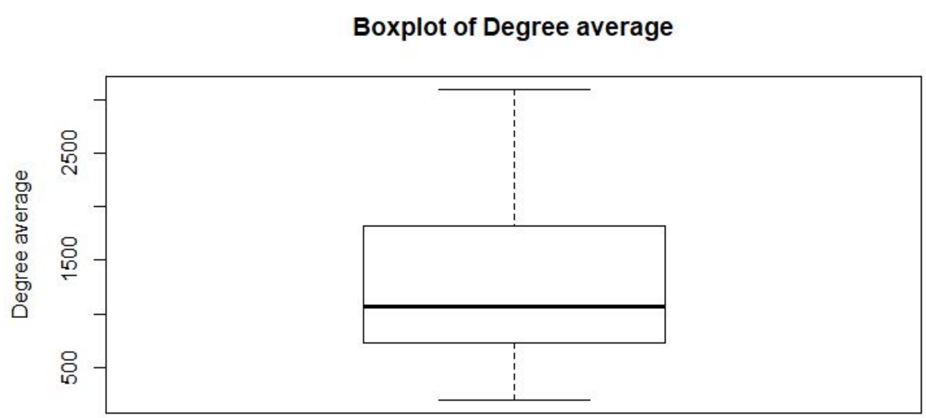
The pictures show the correlation between banks (by considering each bank is a node and the relationships of banks are edge). As we can see with 4 above pictures, there are several statistical figurations like number of nodes, density of graph, diameter and mean distance of graph . The bigger nodes mean that that banks have a higher amount of transferred money. Some biggest nodes are colored by red. We can see that in the last graphs, there are more banks join the bank network, the amount of money transferred is also higher.

3. The evolution of all graphs

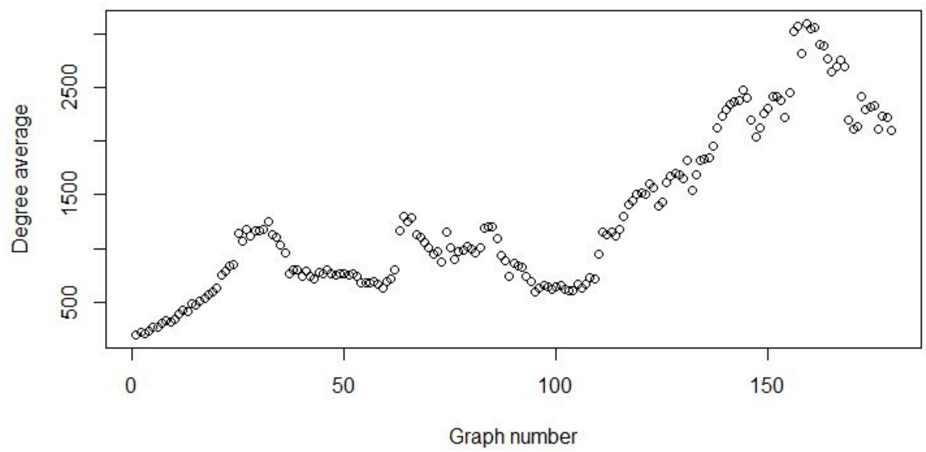
In this section, we collect some key parameters of all 210 networks and make some analysis. In the first case, we start with the degree for nodes in all graphs. But, there are some problems. The first is that there is a wide range number of nodes that lead to hard evaluate of the evolution of all networks. Furthermore, there are too many outliers.



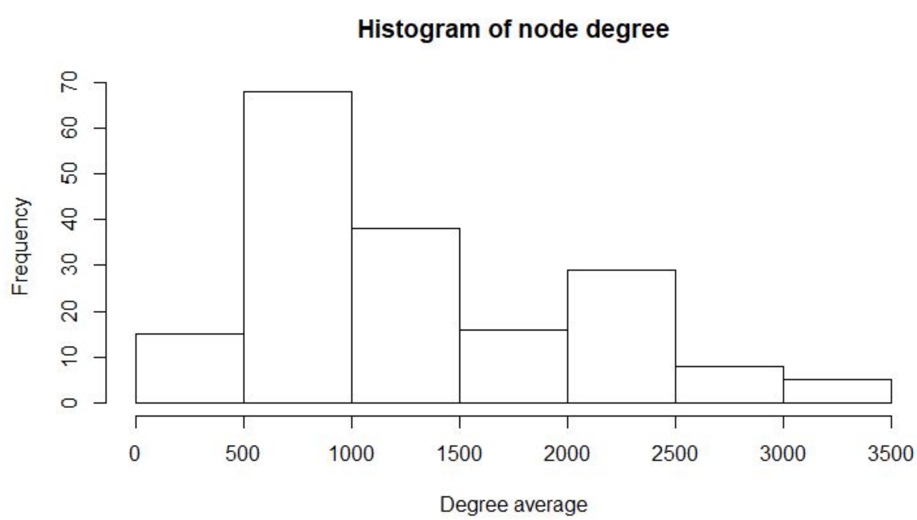
So, we decided to use the Degree average of each graph. The median Degree average is 1000 and 75% of networks has less than 2000.



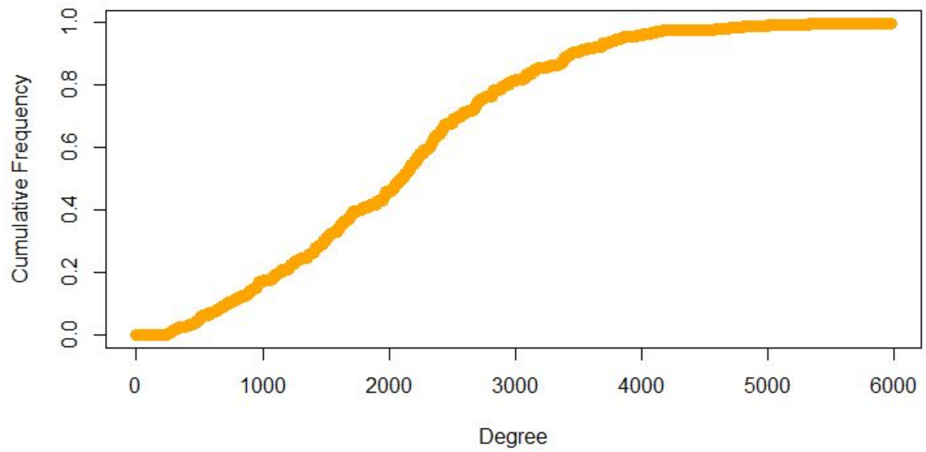
To be more specific, take a look at the scatter plot. The first 100 networks is quite similar to each others. But, there is a clearly upward trend for last networks. In my opinion, it mean that the networks is getting bigger and bigger.



The histogram shows us a different view about all networks. We can see that most of networks has around 1000 or 2000 Degree average. It mean that there are two main model for networks, the simple one and a more complicate one.



The final graph show the cumulative frequency of the degree of all graphs. In this graph, the median is about 1800 and 90% of nodes has less than 4000 degree.



4. Appendix: The code used in this report

```
# Need to install some relevant packages

library(rmatio)

library(igraph)

#This part is to play with some statical analysis in several graph (4 graphs), let check 10th, 40th, 110th and 150th graph in the dataset

for (i in c(10, 40,110,150)) {

 input = paste("/adj_mat_",i,".mat", sep="")

 mat = read.mat(file = input)

 checkna = is.nan(mat$adj)

 chk = mat$adj>0

 adjstar=mat$adj

 adjstar[which(chk)] = 1

 adjstar[which(checkna)] = 0

 net = graph_from_adjacency_matrix(adjstar)

 ########################################

 dg = degree(net, mode="all")

 n=vcount(net)
```

```
  density = ecount(net)/(n*(n-1))

  diameter = diameter(net, directed = TRUE)

  mean_distance = mean_distance(net)

  title = paste("n=", n, "density=",density,"diameter=",diameter,"mean_distance=",mean_distance)

  ######################################

  vcol = rep("grey80", n)

  vcol[unlist(largest_cliques(net))] = "red"

  l = layout_with_fr(net)

plot(net, vertex.label=NA, vertex.color=vcol, cex.main=3.5, main=title, edge.arrow.size=.02, vertex.size=(dg*20/n), edge.curved=0.2, layout=l, rescale=T)

}


# This part is to implement node degree and degree distribution in all graphs

degree_average = c()

degree = c()

for (i in c(1:210)) {

  input = paste("/adj_mat_",i,".mat", sep="")

  mat = read.mat(input)

  adjstar=mat$adj

  net = graph_from_adjacency_matrix(adjstar)

  node_number = vcount(net)

  dg = degree(net, mode="all")

  sum = sum(dg)

  degree_average = append(degree_average,c(sum/node_number))

  degree=append(degree,c(dg))

}


#Display some plots and histograms correspondings to the degree average with a collection 210 graphs
```

```
boxplot(degree, ylab="Degree", main="Boxplot of degree")

boxplot(degree_average, ylab = "Degree average", main = "Boxplot of Degree average")

plot(degree_average, ylab = "Degree average", xlab = "Graph number")

hist(degree_average,main="Histogram of node degree", xlab = "Degree average")


# Degree Distribution

deg.dist <- degree_distribution(net, cumulative=T, mode="all")

plot(x=0:max(dg),   y=1-deg.dist,   pch=19,   cex=1.2,   col="orange",xlab="Degree",   ylab="Cumulative Frequency")
```