CS-E5740 Complex Networks,
Answers to exercise set 6

Sara Cabodi, Student number: 784287

November 4, 2019

Compile with `pdflatex ex_template.tex`

Problem 1

a) Given the 4 different definitions of centrality, here follows their computation for the network of Figure 1.

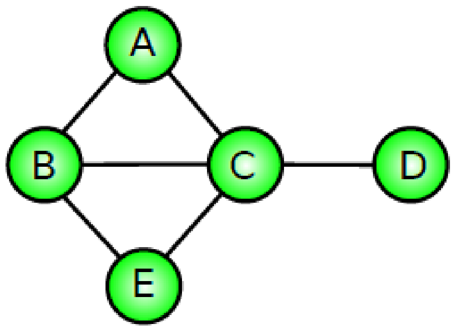


Figure 1: Small undirected network.

- **Degree centrality**: for each node I compute its number of neighbors.

k_A	k_B	k_C	k_D	k_E
2	3	4	1	2

- **Betweenness centrality**: I take every pair of nodes and count how many times a given node belongs to a shortest path between the two nodes of the pair.
  I start by computing σ_{st} (number of shortest paths between s and t) for each pair of nodes s and t (note that this is an undirected network, so (s,t) is equal to (t,s)).

(A,B)	(A,C)	(A,D)	(A,E)	(B,C)	(B,D)	(B,E)	(C,D)	(C,E)	(D,E)
1	1	1	2	1	1	1	1	1	1

Now for each node i I compute the betweeneess centrality with the following formula

$$bc(i) = \frac{1}{(N-1)(N-2)} \sum_{s \neq i} \sum_{t \neq i} \frac{\sigma_{sit}}{\sigma_{st}}$$

where $N = 5$ and σ_{sit} is the number of paths that contain i.
For the nodes A, D and E $bc(i) = 0$, because there aren't any shortest paths passing through them. Regarding B and C, there is for both the path (A,E) that contributes for 0.5 (1 for the single node and 2 is the number of shortest paths from A to E). C is also part of other shortest paths, such as (A,D), (B,D) and (D,E).

$$bc(B) = \frac{1}{(5-1)(5-2)} \left(\frac{1}{2} \right) = \frac{1}{24}$$

$$bc(C) = \frac{1}{(5-1)(5-2)} \left(\frac{1}{2} + 1 + 1 + 1 \right) = \frac{7}{24}$$

- **Closeness centrality**: for each node I compute the inverse of the average shortest path distance to all other nodes than i, applying the formula

$$C(i) = \frac{N-1}{\sum_{v \neq i} d(i,v)}$$

The table of distances is the following

$d(A,B)$	$d(A,C)$	$d(A,D)$	$d(A,E)$	$d(B,C)$	$d(B,D)$	$d(B,E)$	$d(C,D)$	$d(C,E)$	$d(D,E)$
1	1	2	2	1	2	1	1	1	2

$$C(A) = \frac{5-1}{1+1+2+2} = \frac{4}{6} = \frac{2}{3}$$

$$C(B) = \frac{5-1}{1+1+2+1} = \frac{4}{5}$$

$$C(C) = \frac{5-1}{1+1+1+1} = \frac{4}{4} = 1$$

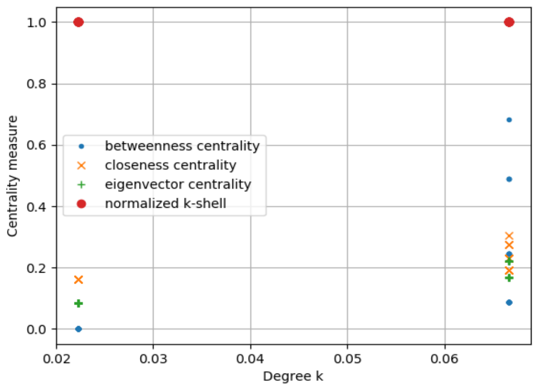
$$C(D) = \frac{5-1}{2+2+1+2} = \frac{4}{7}$$

$$C(E) = \frac{5-1}{2+1+1+2} = \frac{4}{6} = \frac{2}{3}$$

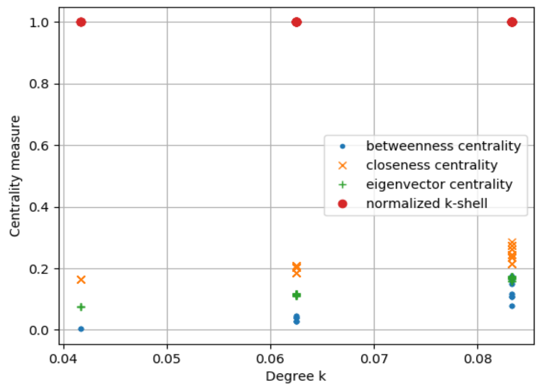
– **k-shell centrality**: the $k_s(i)$ represents the k-shell the node i belongs to. It means that if i belongs to the k-core of the network, but does not belong to the $k+1$-core, than $k_s(i) = k$.

$k_s(A)$	$k_s(B)$	$k_s(C)$	$k_s(D)$	$k_s(E)$
2	2	2	1	2

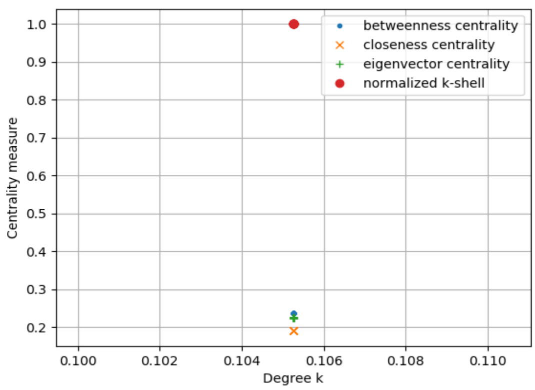
b) The scatterplots obtained from the computations of different centrality measures over the four networks are presented below



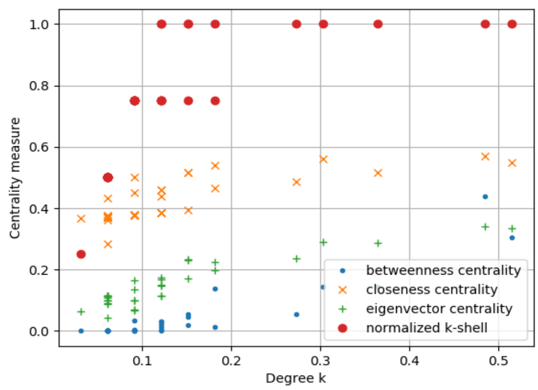
(a) Caley tree network



(b) Lattice network



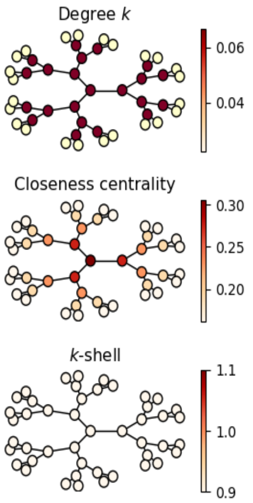
(a) Ring network



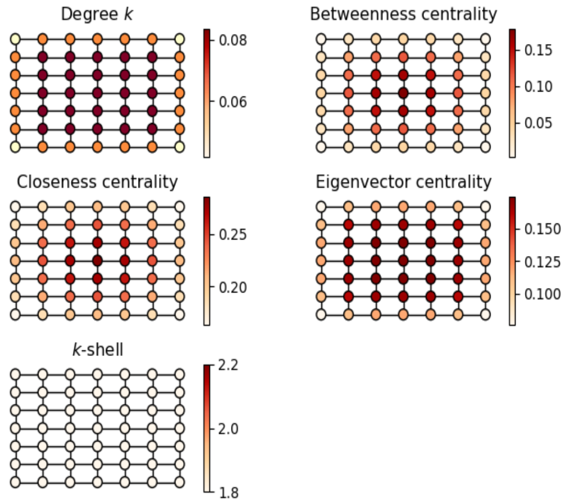
(b) Karate network

c) The network visualizations obtained are reported below

d) Centrality measures represent different ways to score the importance of a given node. Based on definitions in lecture 5 and on the experimental observation of network coloured figures and scattered centrality plots:



(a) Caley tree network



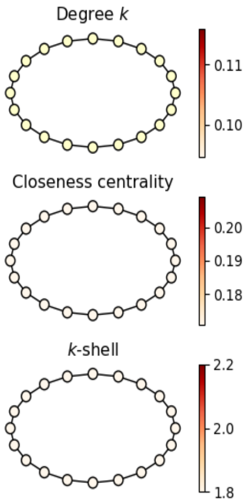
(b) Lattice network

- The ring network is fully balanced, as all nodes have same centrality.
- The lattice and Cayley tree networks provide similar measures: k-shells do not discriminate as, given the network shape, all nodes belong to the same shell. No larger shell exists as peripheral node removal will trigger a wave of other node removals. Degree and eigenvector centrality are similar: they show that peripheral nodes are less important, whereas non peripheral nodes are more important (scoring is a bit more gradual in eigenvector centrality). This seems to indicate that eigenvector centrality is highly related to degrees in this kind of networks. As a matter of fact, both centralities are measures of connectivity. Closeness and betweenness centrality measures show that central nodes are more important. In these two networks, the role of nodes as bridges (beweenness) and the fact of being close to other nodes (closeness) give similar measures.
- The karate network has a less regular shape. In this case, k-shell clearly shows that the network has a core central subnetwork, whose nodes have high closeness and eigenvector centrality. Degree and betweenness scores provide less hints in order to discriminate node centrality, as most nodes have low scores.

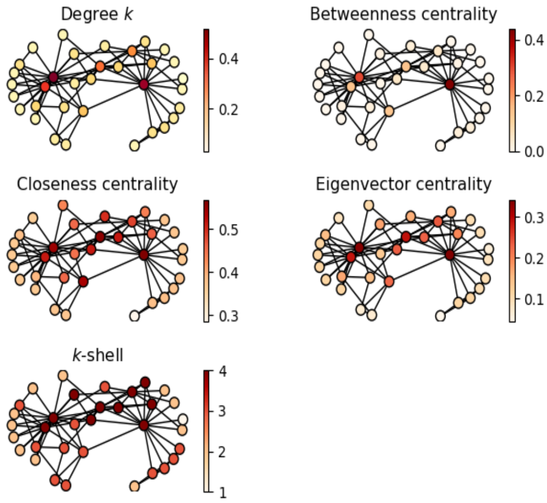
Problem 2

a) The two scatter plots of the degrees of pairs connected nodes for the two networks (karate club and Facebook) are reported below

b) The two heat maps of the degrees of all connected nodes for the two networks (karate club and Facebook) are reported below



(a) Ring network



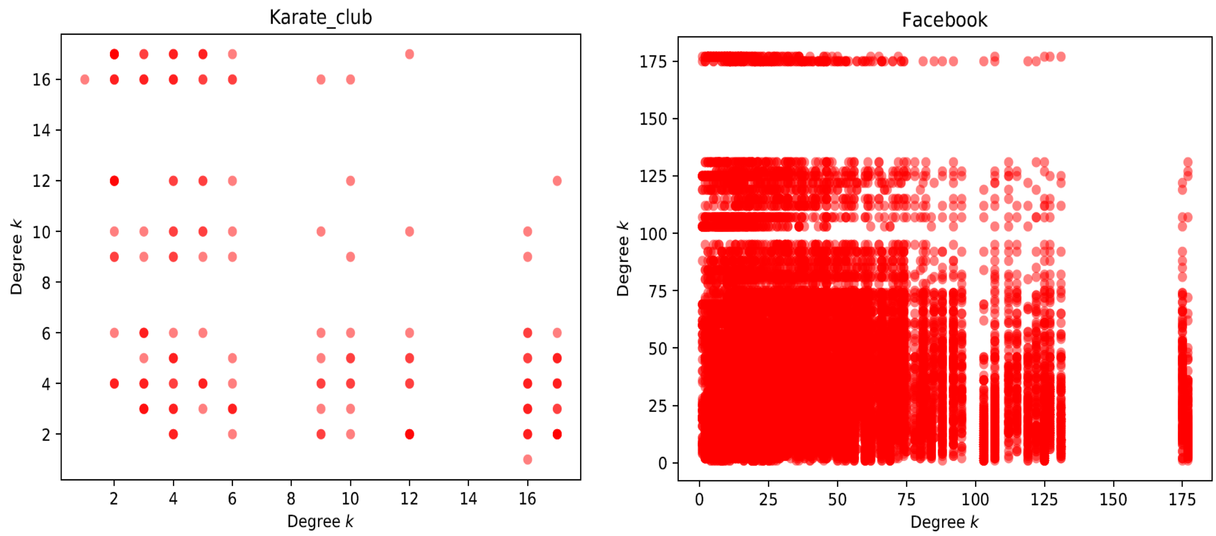
(b) Karate network

The scatter plot tells us only if, given two degrees k_i and k_j, they are connected. The heat map adds information regarding how much they are connected (how many connected pair of nodes with degrees k_i and k_j).

c) The assortativity coefficient is defined as Pearson correlation coefficient of the degrees of pairs of connected nodes.
The assortativity measures are:

- Karate club network = -0.4756
- Facebook network = 0.05598

The Karate Club network is a well known example of a small social network generated from a university club. The network shows two main clusters: it is disassortative, as the cluster hubs are not directly connected, but each one (node with high degree) is connected with nodes with low degree. This is confirmed by the negative (close to -0.5) assortativity coefficient. Furthermore, by analyzing the scatter plots and heat map of degrees, it seems that edges are rather uniformly distributed on degrees. The heat map also shows that high degree nodes are mainly connected with low degree nodes, which confirms the result on assortativity. On the other hand, Facebook (a popular and widespread) is expected to be assortative. The near 0 value partly contradicts this assumption. A possible explanation is that Facebook is undoubtedly including assortative nodes and sub-networks/clusters, but also parts that are not assortative, so on average we obtain an intermediate (neutral) result. The scatter plot and heat map of degrees suggest that most direct connectivity is between couples of nodes with lower degree (up to degree 25 or 30). The scatter plot also shows a clear distinction between a subset of high degree nodes (k of about 175, that have no



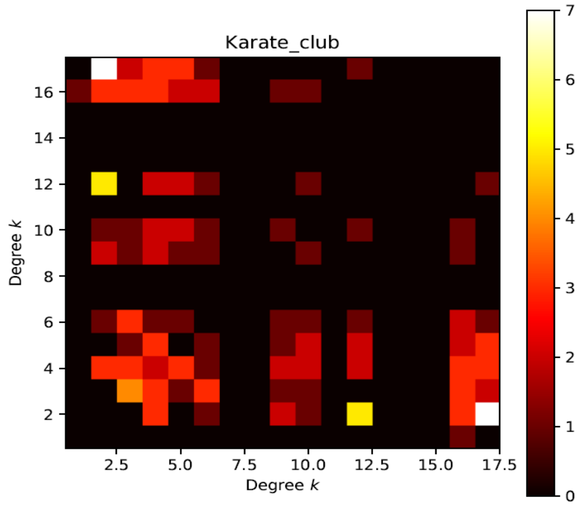
(a) Karate club network

(b) Facebook network

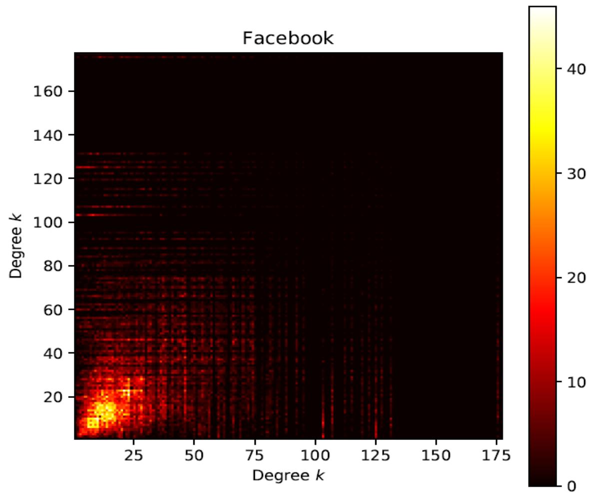
direct connection each other) and all other nodes with k below 130. So the $k = 175$ hubs are not mutually reachable, whereas hubs with $k = 125$ (or close values) are mutually connected, which confirms the intermediate assortativity value.

d) The two scatter plots of k_{nn} as function of degree k are presented below

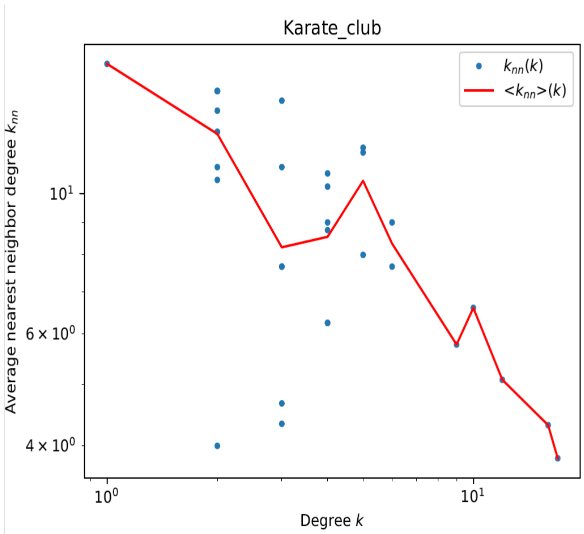
The result for the Karate club network agrees with the one of assortitivity. Since the $\langle k_{nn} \rangle (k)$ function is decreasing, the network is disassortative (negative assortativity). As for the Facebook result, the trend of the $\langle k_{nn} \rangle (k)$ function neither increasing nor decreasing highlights the fact that it is a weak assortative network, because it has a positive assortativity, but very low in module.
Both scatter plots of k_{nn} show a distribution of values that is much wider at low k values. For high k degrees, k_{nn} values are very close to the their average. A possible explanation is that the (few) high degree nodes behave in a uniform way (low assertativity, which means connected to low degree nodes, in the Karate network and high/average assertativity for Facebook). Nodes with low degree have a wider range of connectivity, as they are connected to nodes ranging from low to very high degree.



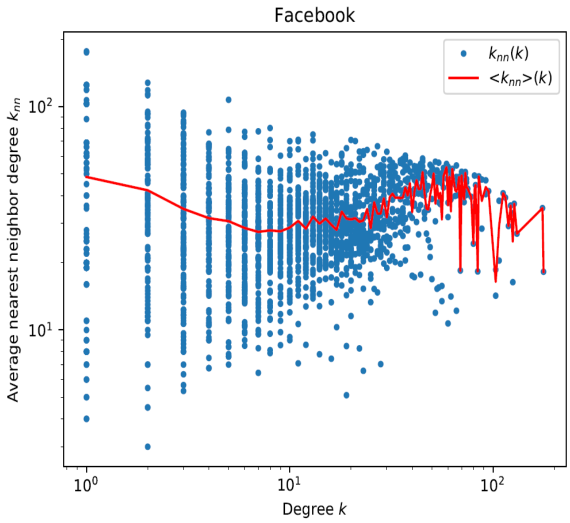
(a) Karate club network



(b) Facebook network



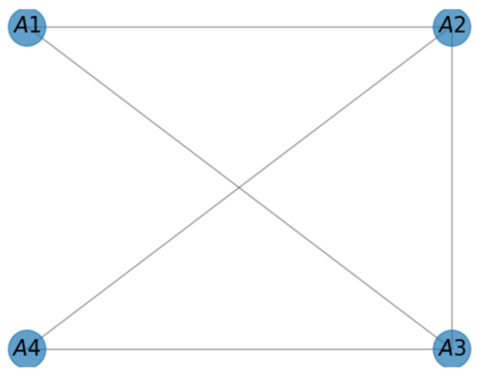
(a) Karate club network



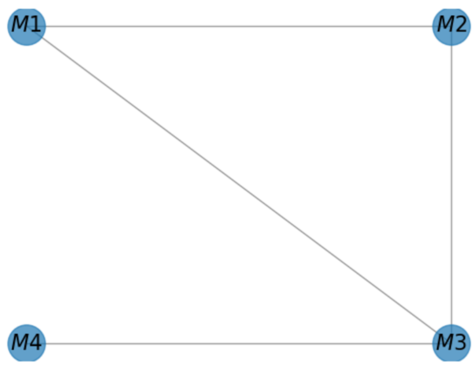
(b) Facebook network

Problem 3

a) Here are the two unipartite projections of the network, one for the actors and one for the movies.



(a) Actors network



(b) Movies network

b) Here is another bipartite network with the same unipartite projections as the one shown in the exercise sheet in Figure 3

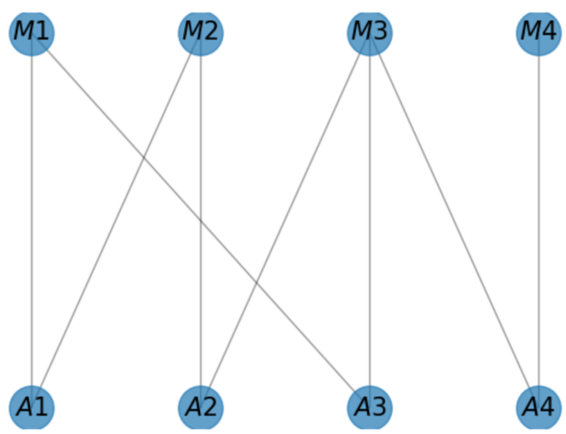


Figure 10: Bipartite network of actors (A) and movies (M).