CS-E5740 Complex Networks, Answers to exercise set 6

Yangzhe Kong, Student number: 765756

November 2, 2019

Problem 1

a) 1) Degree Degree $k(i)$ is the number of neighbors of node i.

Node	Degree
A	2
B	3
C	4
D	1
E	2

2) Betweenness centrality $bc(i)$ Number of shortest paths between other nodes of the network that pass through node i. Formally, it is

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

In our case $N = 5$. And node A, D, E doesn't lie on any shortest paths of other nodes. Node B lies on the shortest path $A \to B \to E$.
Node C lies on the shortest paths $A \to C \to E$, $A \to C \to D$, $B \to C \to D$, $D \to C \to A$, $D \to C \to B$, $D \to C \to E$, $E \to C \to A$, $E \to C \to D$.
Thus we have the following betweenness centrality for each node:

Node	Betweenness centrality
A	0
B	$\frac{1}{12} = \frac{1}{12} \times (\frac{1}{2} + \frac{1}{2})$
C	$\frac{7}{12} = \frac{1}{12} \times (6 + \frac{1}{2} + \frac{1}{2})$
D	0
E	0

3) closeness centrality $C(i)$ Inverse of the average shortest path distance to all other nodes than i.

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

Then we can calculate the sum of shortest path for each node.
For node A, sum of shortest path is 6. For node B, sum of shortest path is 5. For node C, sum of shortest path is 4. For node D, sum of shortest path is 7. For node E, sum of shortest path is 6.

Node	Closeness centrality
A	2/3
B	4/5
C	1
D	4/7
E	2/3

4) k-shell $k_s(i)$
Node i belongs to the k-shell, if it belongs to the k-core of the network but does not belong to the $k+1$-core. The k-core is the maximal subnetwork (i.e. the largest possible subset of the network's nodes, and the links between them) where all nodes have at least degree k. In other words, the 1-core is formed by removing nodes of degree 0 (isolated nodes) from the network, the 2-core is formed by removing nodes of degree 1 and iteratively removing the nodes that become degree 1 or 0 because of the removal, the 3-core is formed by removing nodes of degree less than 3 and those nodes that become of degree less than 3 because of the removals, and so on. The 1-shell is then the set of nodes that was removed from the 1-core to obtain the 2-core.
So then we have to form the k-cores first. For 1-core, the set of nodes is $\{A, B, C, D, E\}$. By removing nodes of Degree $k(i) = 1$, we can get the set of nodes for 2-core is $\{A, B, C, E\}$. And there is nothing in 3-core and so on.

i-core	Set of nodes
1-core	$\{A, B, C, D, E\}$
2-core	$\{A, B, C, E\}$

b) visualization of betweenness, closeness, k-shell, and eigenvector centrality as a functionof degree in a scatter plot for each of the networks is shown in Figure 1, 2, 3 and 4.

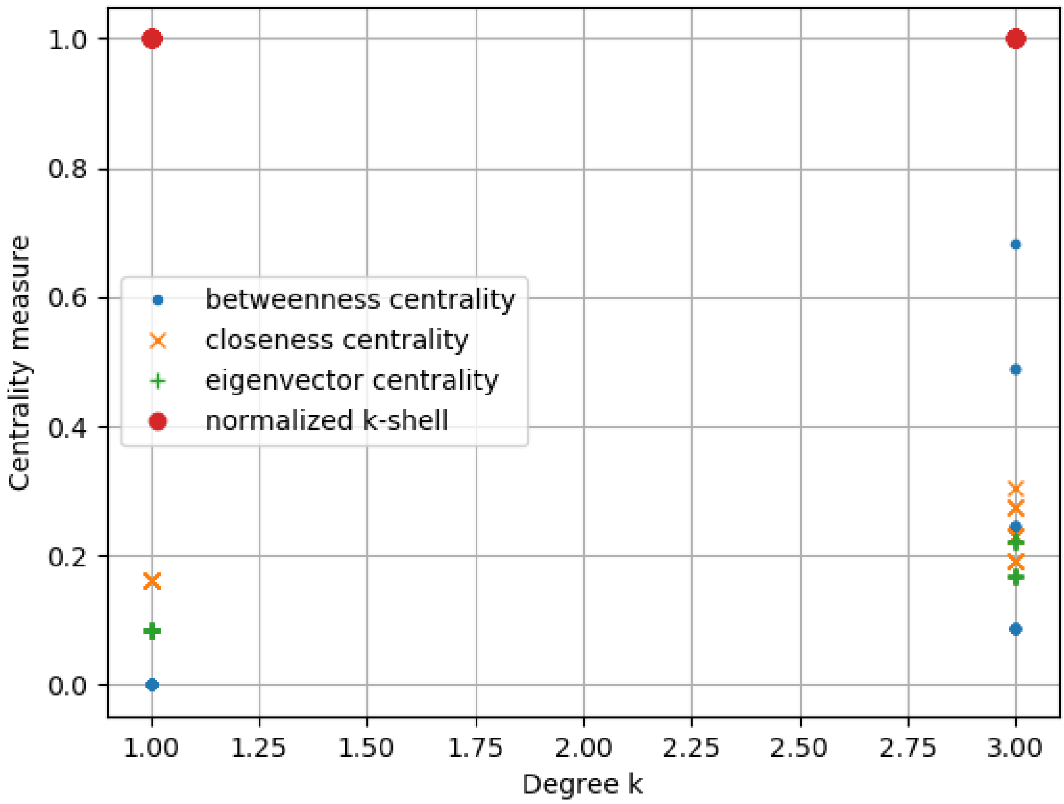

Figure 1: Centrality Measures of Cayley Tree

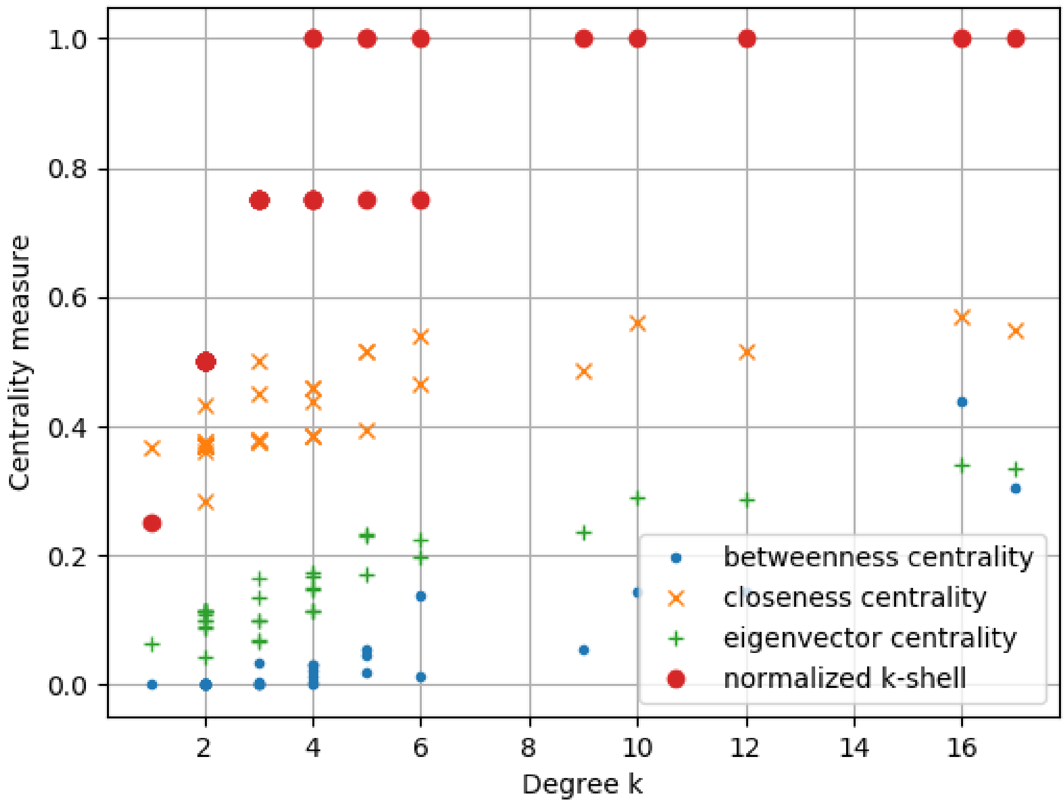

Figure 2: Centrality Measures of Karate Club Network

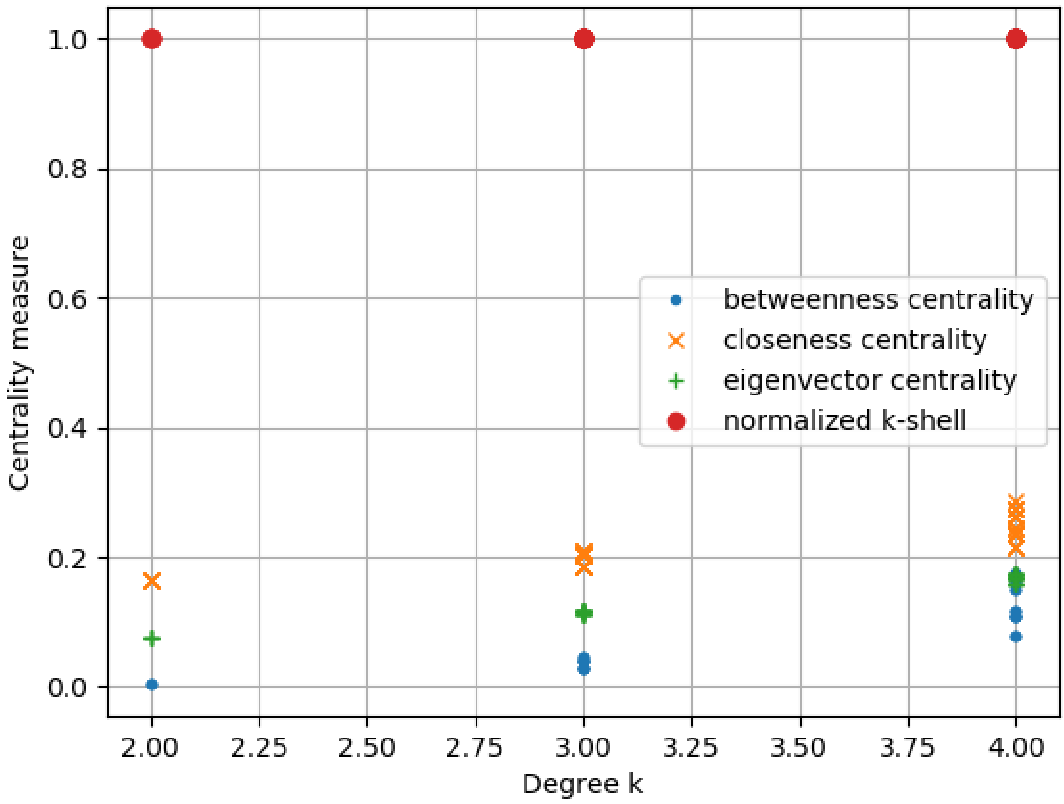

Figure 3: Centrality Measures of Lattice Network

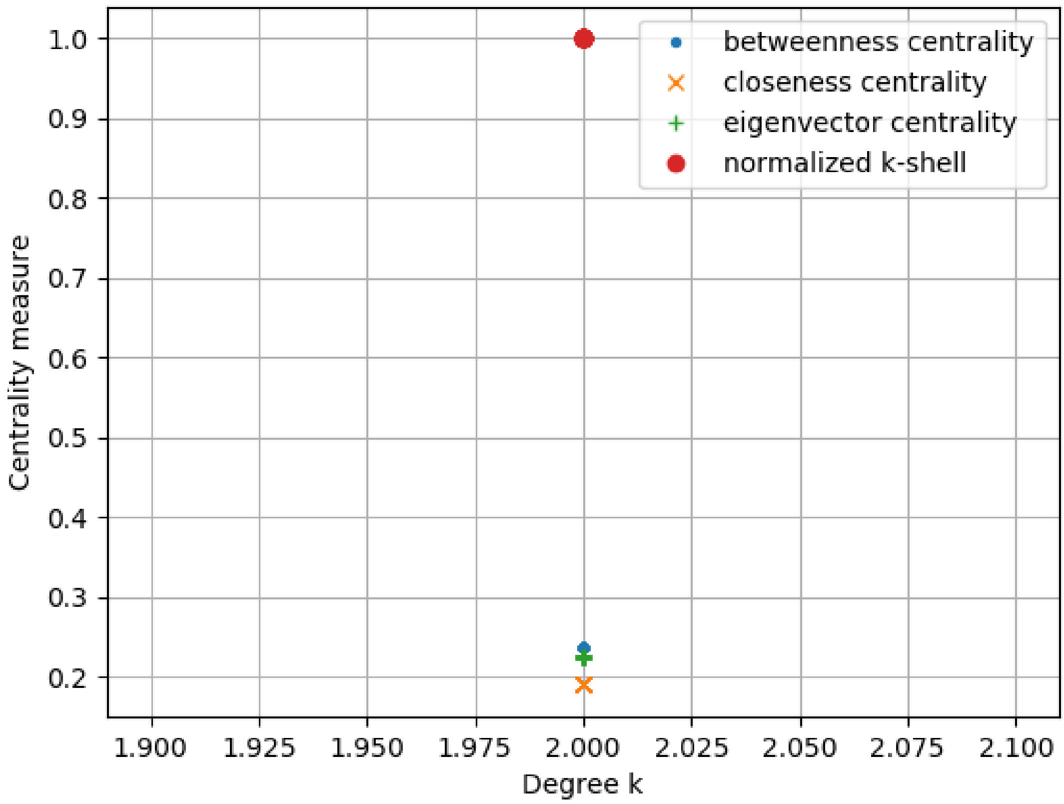

Figure 4: Centrality Measures of Ring Network

c) Visualizations of all the networks are shown in Figure 5, 6, 7 and 8 using one of the centrality measures each time to define the colors of the network nodes.

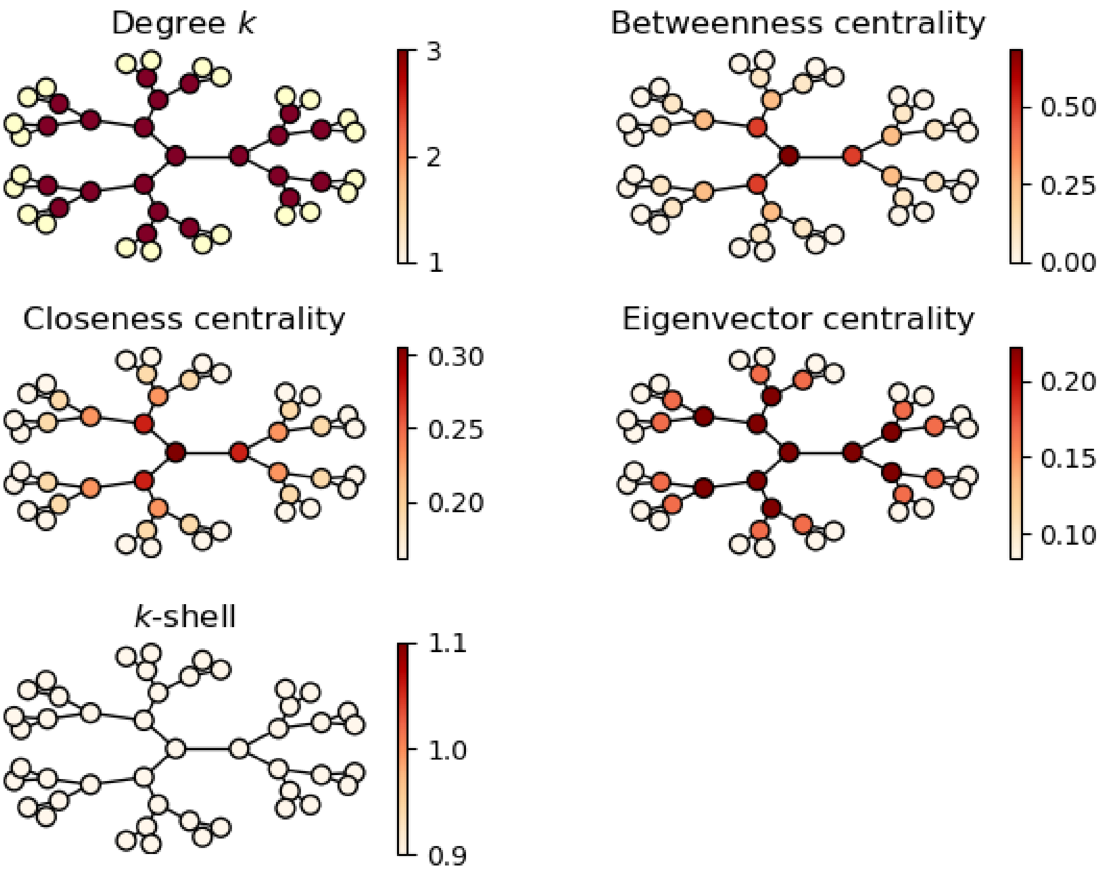

Figure 5: Visualization of Cayley Tree

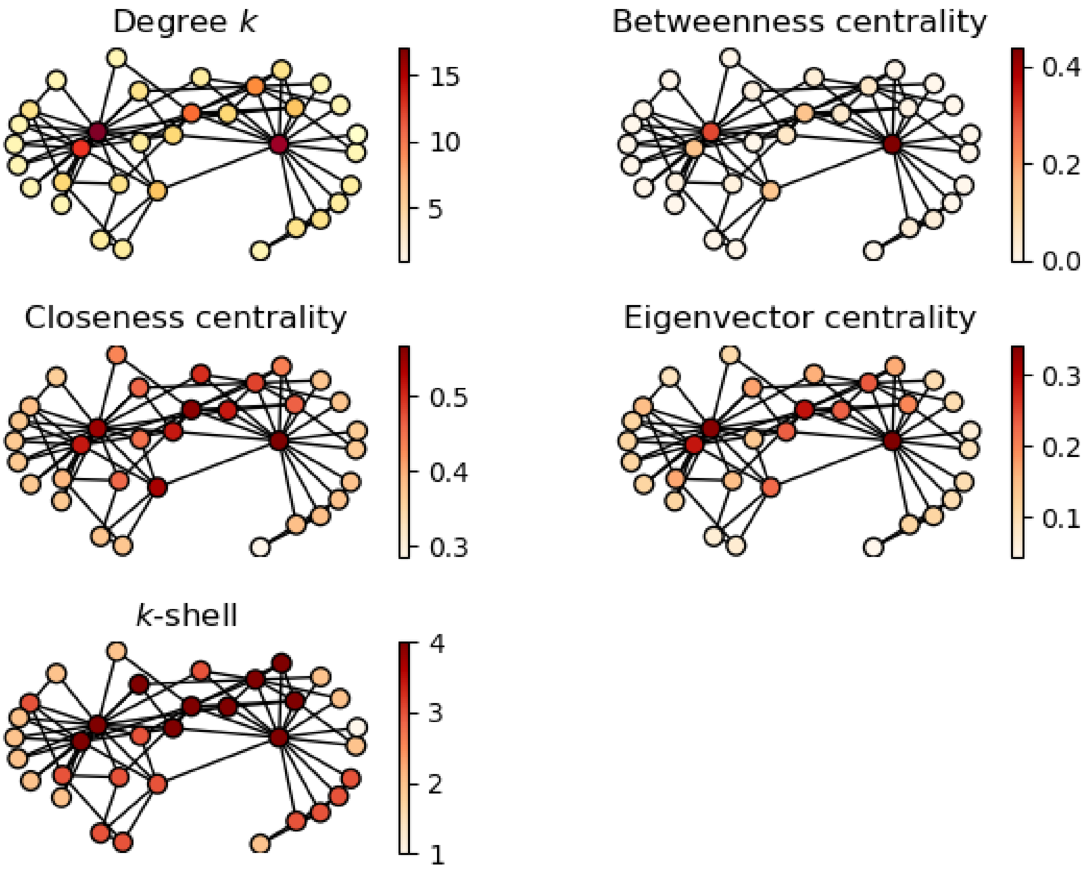

Figure 6: Visualization of Karate Club Network

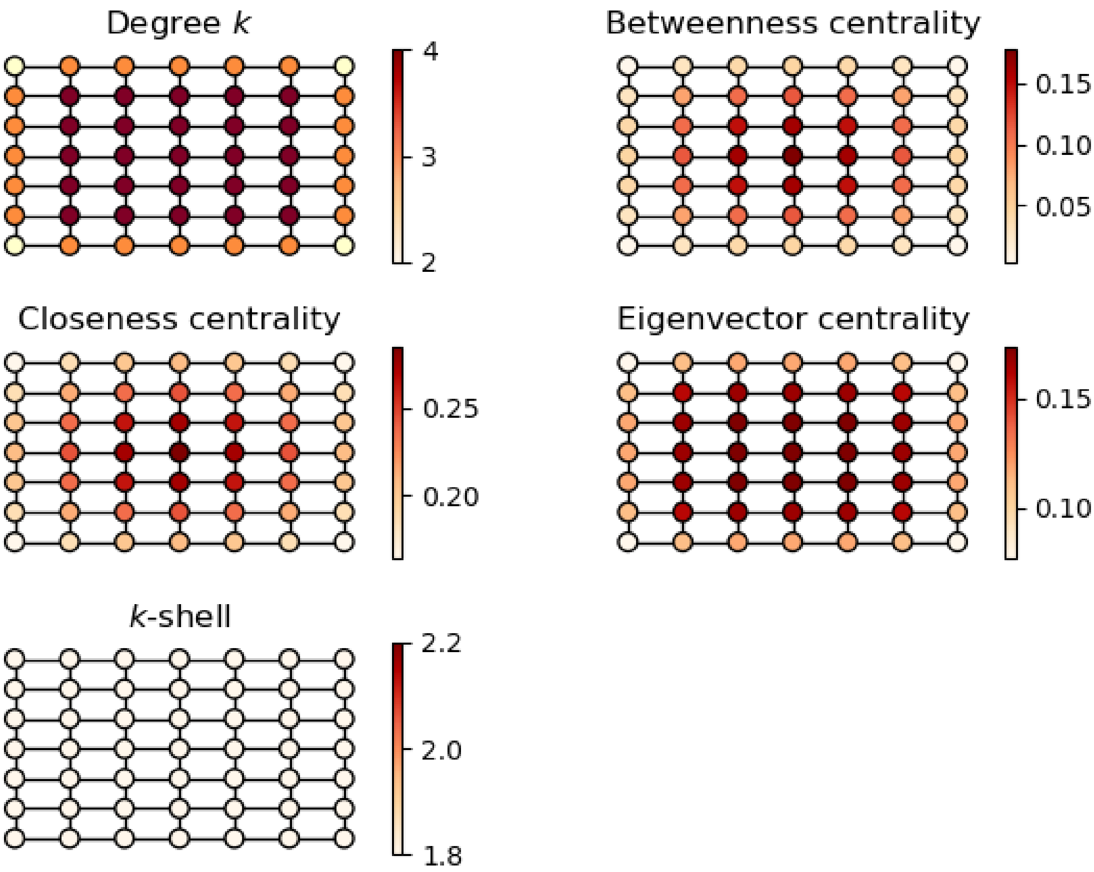

Figure 7: Visualization of Lattice Network

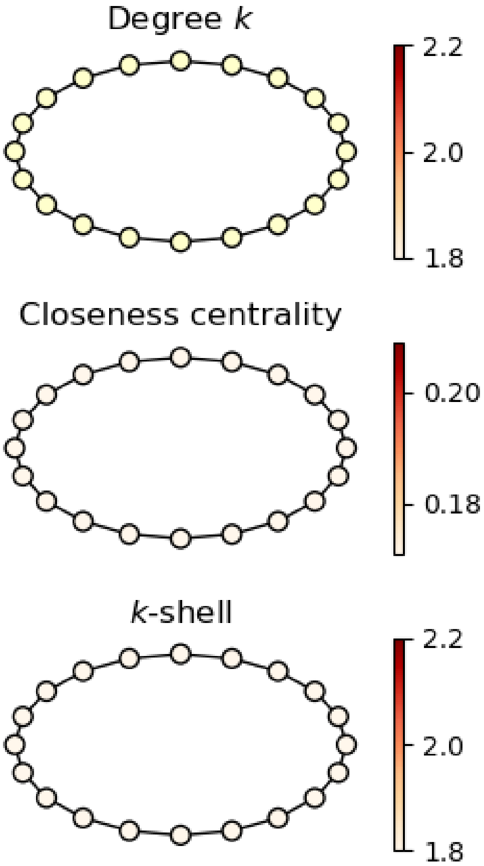

Figure 8: Visualization of Ring Network

d) The k-shell centrality of nodes only have significant differences in Karate Club network. It measures how similar two well-connected nodes are.
The betweenness centrality depends on the shortest paths. And closeness centrality depends on the distance of a node to other nodes. It's pretty natural to observe that these two centrality measures give quite similar results, as observed in the 2d-lattice and cayley trees. These 2 centrality measures give higher values to nodes in the centre (which we can think of being more important) and lower values to the nodes that are further away from the centre.
The eigenvactor centrality depends on both the degree of the node we concern as well as its neighbors. This centrality seems to perform somewhere between the degree centrality and the betweeness/ closeness.

Problem 2

a) Scatter plots of Karate club network and Facebook friendships network are shown in Figure 9 and 10.

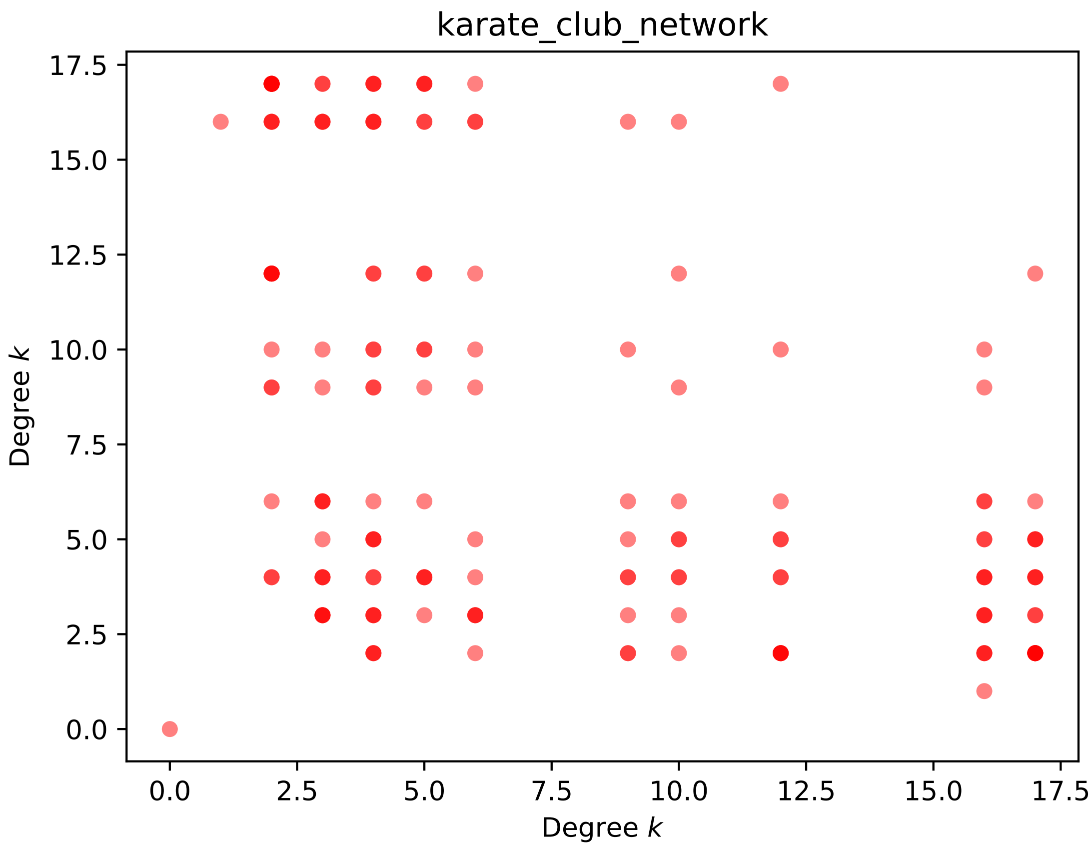

Figure 9: Scatter plot of Karate club network

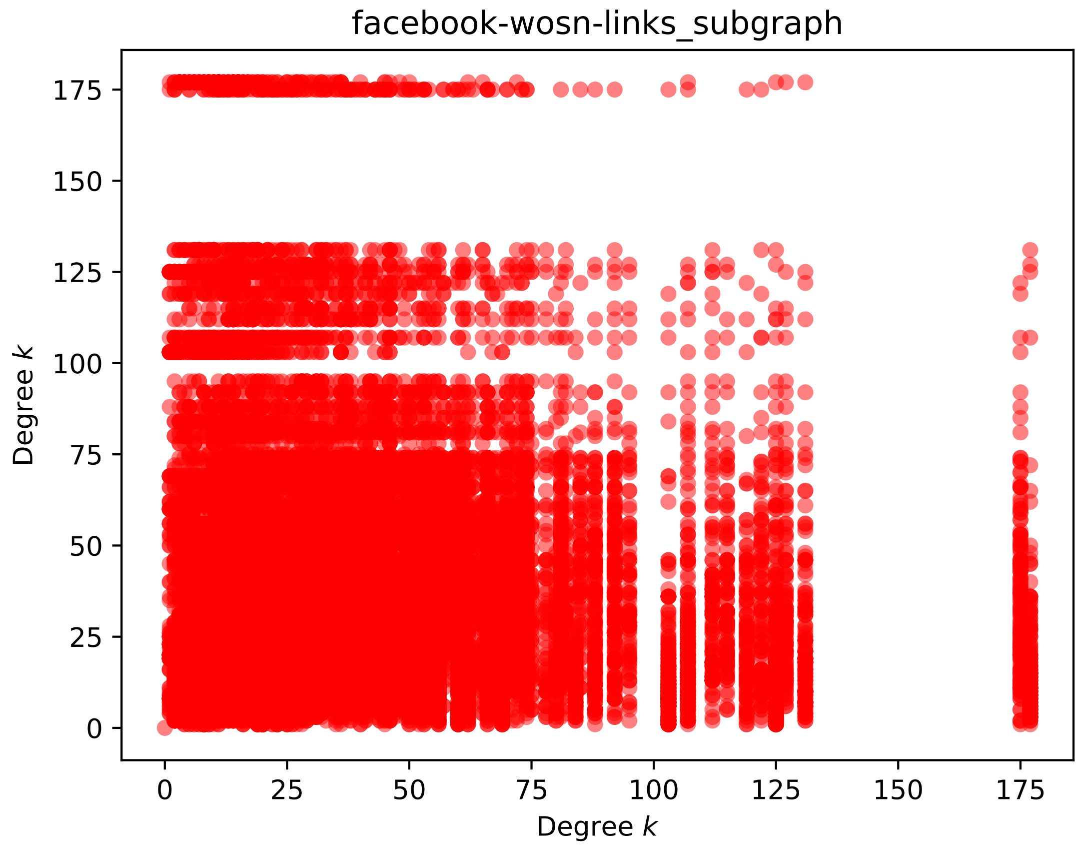

Figure 10: Scatter plots of Facebook friendships network

b) Heat maps of Karate club network and Facebook friendships network are shown in Figure 11 and 12.

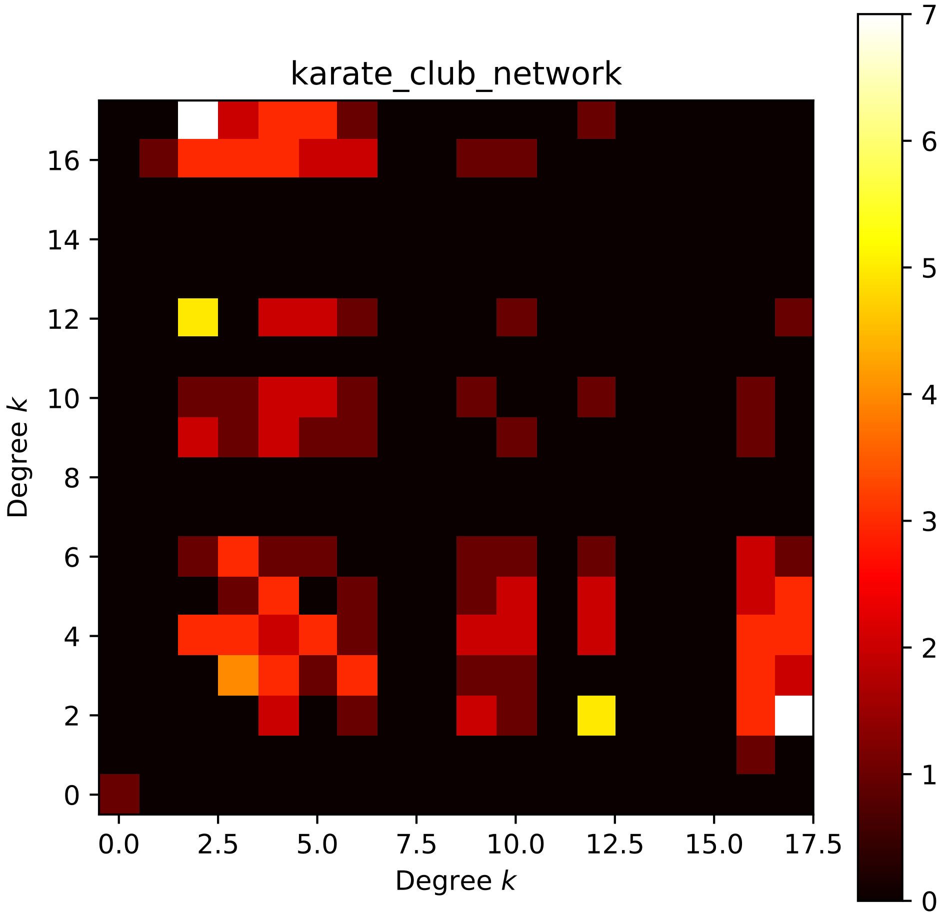

Figure 11: Heat map of Karate club network

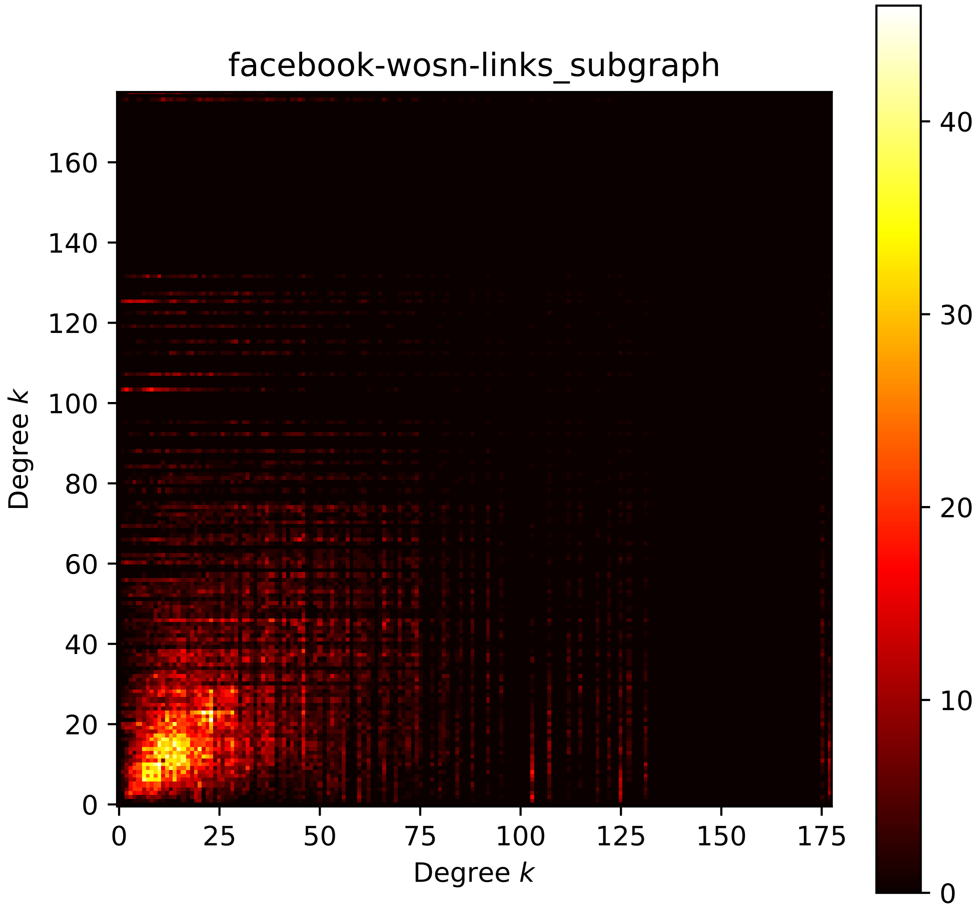

Figure 12: Heat map of Facebook friendships network

We can observe from the figures that the heat map provides 1 more dimension of data: the frequency of the pair of nodes through coloring the bins.

c) Own assortativity for *karate_club_network*: -0.4632586928363579
NetworkX assortativity for *karate_club_network*: -0.47561309768461457
Own assortativity for $facebook - wosn - links\_subgraph$: 0.0559899301505 09945
NetworkX assortativity for $facebook - wosn - links\_subgraph$: 0.05598478476593048

The Facebook friendship network is assortative since it is a social network, but Karate club network is not since the conflicts in the club leads to the split of the club and the forming up of 2 clusters inside the club.

d) Scatter plots of K_{nn} as a function of k of Karate club network and Facebook friendships network are shown in Figure 13 and 14.

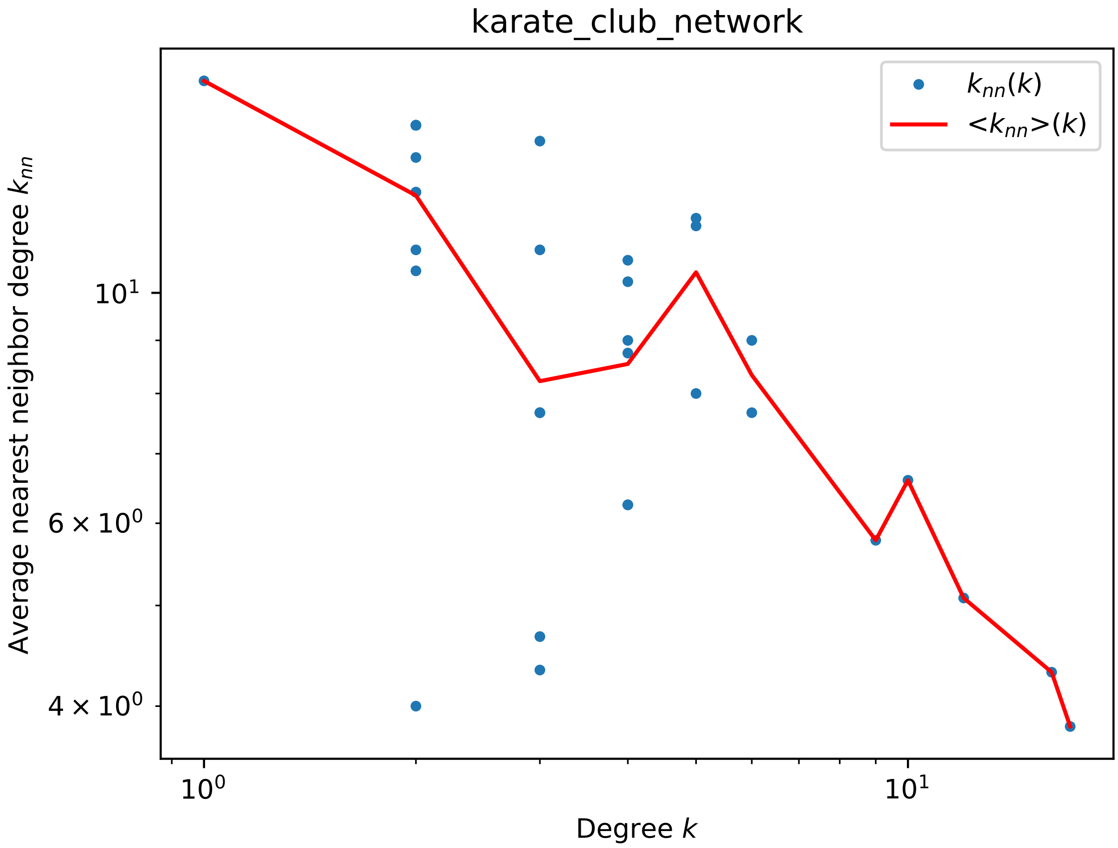

Figure 13: Scatter plots of K_{nn} as a function of k of Karate club network

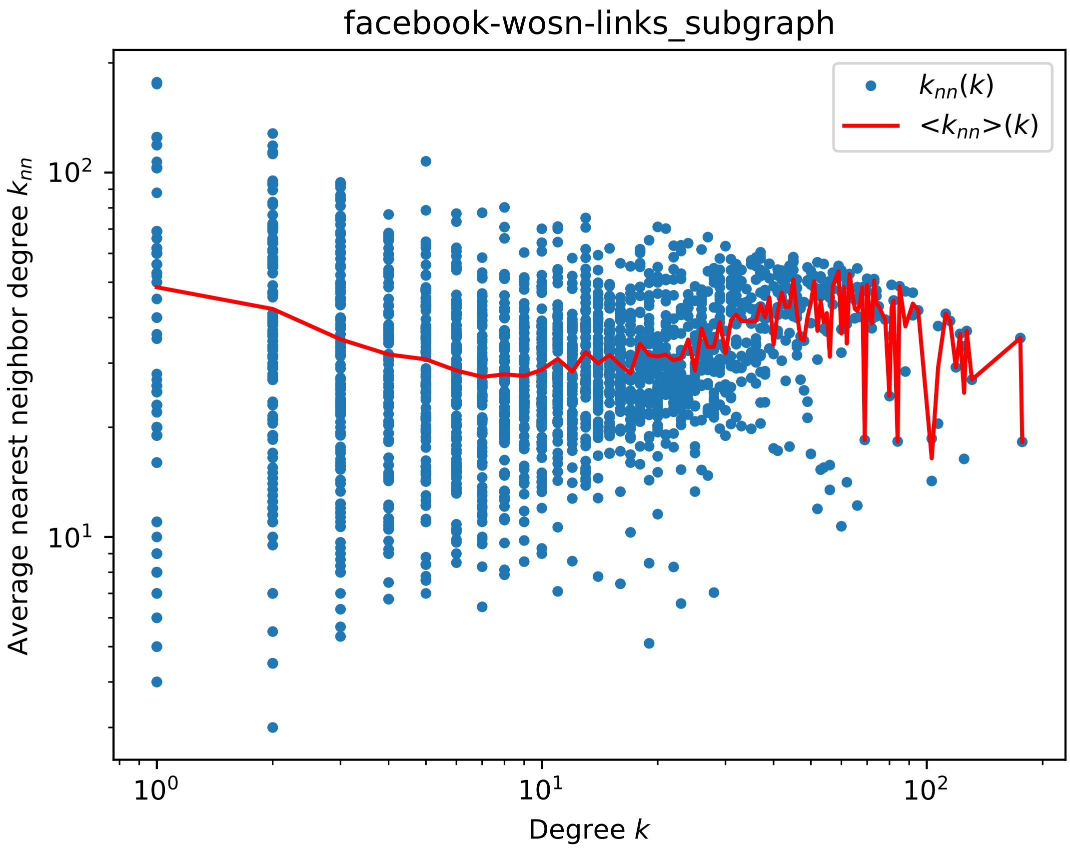

Figure 14: Scatter plots of K_{nn} as a function of k of Facebook friendships network

We can observe from the average neighbor degree distribution given different degree values that for a large social network such as Facebook friendships network, a node with large degree will have higher probability of having also a large average neighbor degree compared with Karate Club Network on contrary. This is mainly because of homophily, since people usually would like to make friends with others just like them. Also since assortativity measures the similarity of connections in the network with respect to the node degree, we can see that in networks with negative assortativity, $\langle k_{nn} \rangle$ would vary widely but in networks with positive assortativity, $\langle k_{nn} \rangle$ won't vary too much.

Problem 3

a) Unipartite projections of the network of actors and the network of movies are shown in Figure 15 and 16.

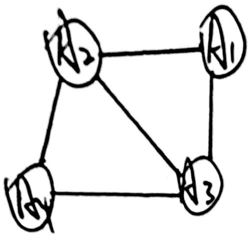

Figure 15: Unipartite projection of the network of actors

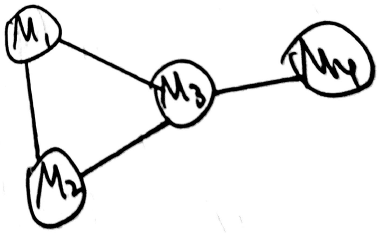

Figure 16: Unipartite projection of the network of actors

b) A counterexample in shown in Figure 17.

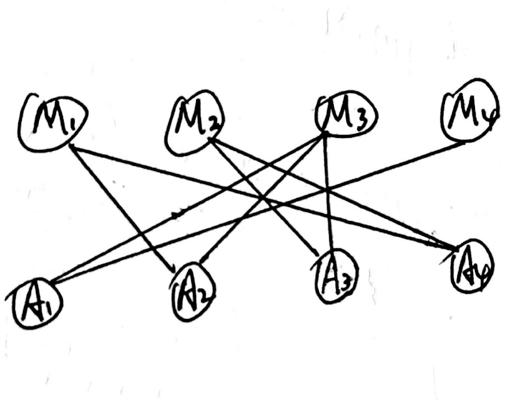

Figure 17: A counterexample