BT5420 - COMPUTATIONAL SYSTEMS BIOLOGY

Assignment 2

Submitted by Sahana (BE17B038)

Question 1

http://networkrepository.com/ENZYMES8.php - This is a Labelled network of enzymes.

The network does not have edge weights. It is an undirected graph.

Assortativity –
This was computed based on two methods, both of which are reported.

The first method is based on the formula given in Textbook (NewMan formula) where the nodes of the network was classified into two categories – high degree and low degree. The threshold was set to be the average degree of the network. A 2*2 matrix was constructed based on the number of interactions between the two types of nodes (This will yield a symmetric matrix).

$$r = \frac{\text{Tr}\ \mathbf{e} - \|\mathbf{e}^2\|}{1 - \|\mathbf{e}^2\|}$$

The assortativity coefficient calculated by this method was found to be 1.3822.

The second method was to get a simple Pearson Correlation coefficient between the left_index_degrees and right_index_degrees of all edges. This was computed as follows:

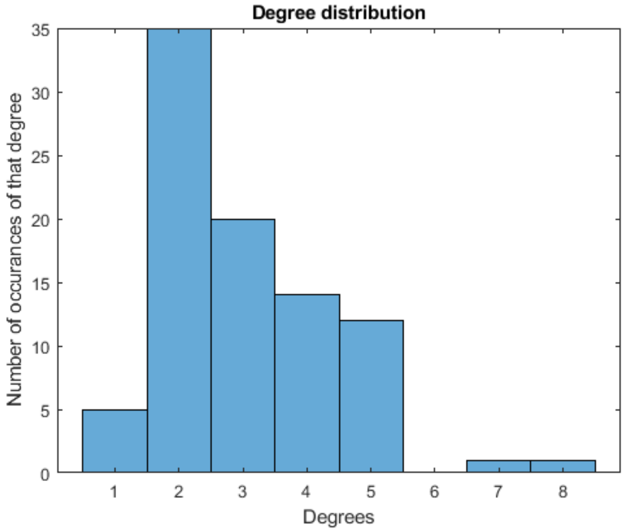
```
as = corrcoef(LR,RL);
```

This will yield a 2*2 matrix with diagonal elements as 1.0 always. The off-diagonal element specifies the R value which is the correlation coefficient between the two vectors.

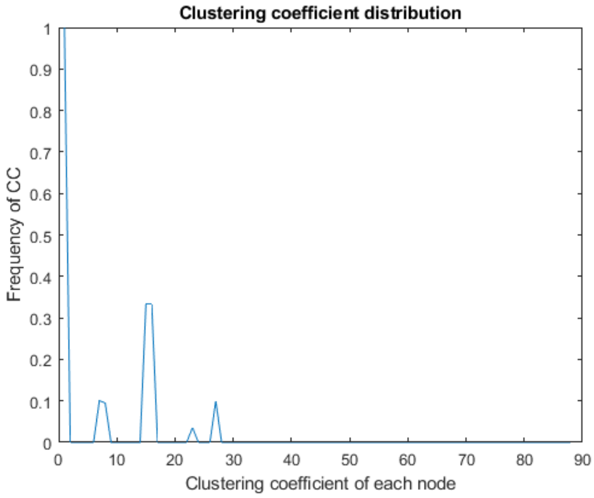
The assortativity coefficient calculated by this method was found to be 0.0685.

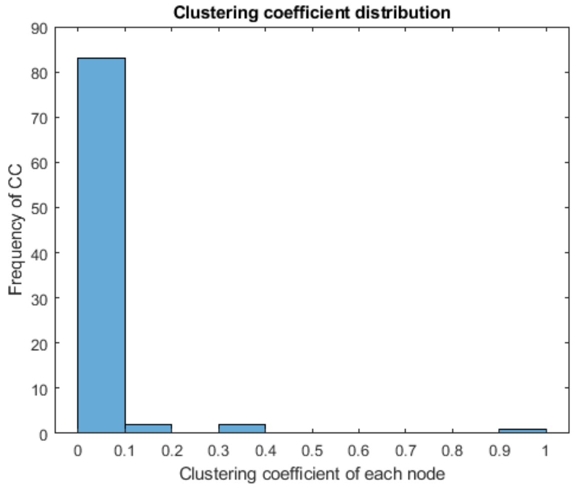
Inference - Therefore, we can conclude that the degrees of the network are very minimally positively correlated/assorted.

Degree Distribution



Distribution of Clustering Coefficient -





(d)

- The network appears to have a degree distribution that follows Power Law.
- The trend of the distribution of Clustering coefficient decreases as the node degree increases. This distribution also follows a power law. This implies that the low-degree nodes belong to very dense sub-graphs and those sub-graphs are connected to each other through hubs.

Both the above characteristics hint towards the same fact that the network resembles a scale-free network.

Question 2

Network A was constructed as a regular lattice with number of nodes = 100 and each node was connected to 3 nodes on the right and 3 to its left (There is no directionality involved. Left and right were mentioned to understand the architecture of the network.)

Rewiring a network -

Network A was rewired in the following way.

I did an iteration through all edges. While we are under the specified probability, I generate a random number. I check if the random number is not the same as neighbours of the left_index of a particular edge, or the left_index itself. In that case, I cut the already existing edge and create a new one between (left index, the generated random number). (Cut and create was done by just changing the right_index of the edge). At the end f every iteration, I updated the network with the changes in edge list (eA as in the program), so that in each round, neighbour vector is found appropriately.

The program is written as follows:

```
function [B] = hw2_2_rewiring(A,p)
```

where A is the network itself (a Graph variable) and p is the probability of rewiring an edge. The function returns B, which is an adjacency matrix of the newly formed/rewired network.

As the rewiring function is called within the main program of 2nd question, the probability of rewiring (p) was input as an argument to this function while calling.

The program is written as follows:

```
function [] = hw2_2(p)
```

Final values – (for input parameter p = 0.2)

Diameter – This is the length of the longest geodesic (longest shortest path!) in the graph.

Network A = 17

Network B = 6

Characteristic path length – This measures the average separation between two nodes in a network.

Network A = 8.7576

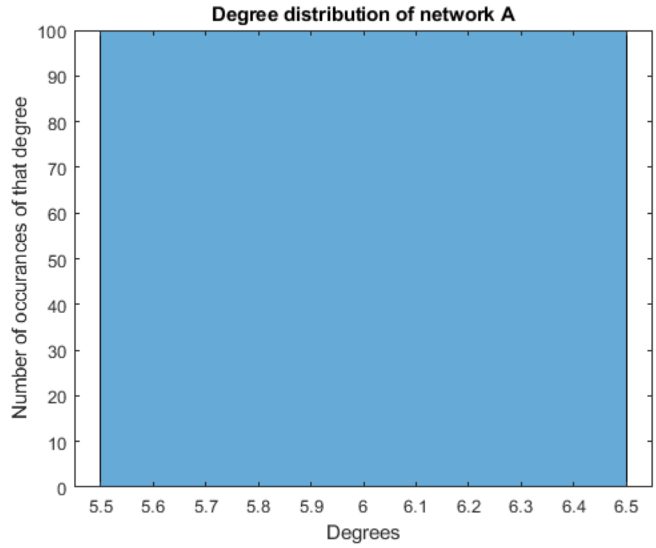
Network B = 3.2133

Global clustering coefficient – C can be thought to measure the cliquishness of a node's neighbourhood. The mean clustering coefficient (CC) of the network is its global CC as well.

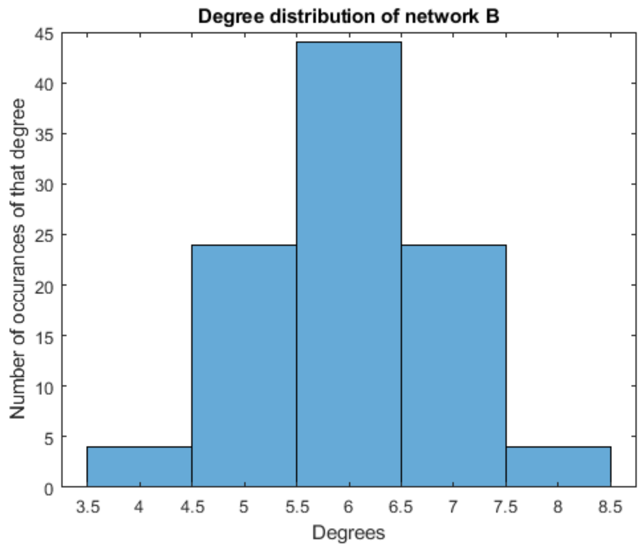
Network A = 0.6000

Network B = 0.3385

Degree distribution of Network A –



Degree distribution of Network B –



Network A and B are as follows –