Assignment 2

BT5240 Computational Systems Biology | Jan - May 2022
Network Biology

February 10, 2022

Due Date: 18th February, 2022
Maximum marks: 25

Academic Integrity: You are allowed to discuss the problems verbally with your friends, but copying or looking at codes (either from your friend or the Web) is not permitted. Transgressions are easy to find, and will be reported to the "Sub-committee for the Discipline and Welfare of Students" and will be dealt with very strictly. Mention any collaboration (discussions only!) in your solutions.

Submission: Since this is a computational assignment, we would also like to look at your codes. Submit your assignment as one zip file by uploading it at http://tinyurl.com/bt5240-submit
Your zip file should be named something like BTyyBxxx.zip, based on your roll number. This zip file must contain a single neatly typeset PDF of your solutions (named BTyyBxxx.pdf) including well annotated the plots and figures of legible fontsize as well as the codes used for each of problems in a separate folder codes with proper annotations.

Problem 1: Characterizing the networks
Choose a real-world network from http://networkrepository.com/ and perform the following analysis. Do mention which network have you chosen and attach the MAT file of the network.

1. Scale-free properties of the network. (10 points)
    a. Plot Degree Distribution (N(k) vs k). (2 points)
    b. Plot the log (N(k)) vs log(k). Compute the spearman correlation coefficient. Report the p-value as well. (3 points)
    c. Fit a linear relationship between log (N(k)) and log (k). Report the goodness of the fit (R^2 $score$) and the coefficients. (3 points)
    d. Comment on the structure of your network. Does the network follow power law? If yes, what is the corresponding γ? (2 points)
2. Small world properties of the network. (15 points)
    a. Compute the average shortest path length of your network (L) and plot the distribution of all the shortest path lengths of your network. (3 points)
    b. Generate 1000 Erdos-Renyi networks of the same size (same number of nodes and edges) and compute the average shortest path length of each of these networks. The mean of these average shortest path lengths would be the 'mean average shortest path length' (L_r) of a random network. (4 points)

c. Plot the clustering coefficient distribution (C(k) vs k) and Compute the average clustering coefficient C of your network. (3 points)
d. Obtain the average clustering coefficient of a perfect lattice of same number of nodes and edges (C_L). (2 points)
e. Compute omega (ω) the 'small worldness' of your network. The formula is given below. (1 point)

$$\omega = \frac{L_r}{L} - \frac{C}{C_L}$$

f. Based on omega and the degree distribution, comment on the structure of the network especially in light of small-world properties of the network. You can refer to Telesford et al. Brain connectivity 2011 to learn more about omega. (2 points) [DOI: 10.1089/brain.2011.0038]