Network Science

Assignment 1

- *Code for the following tasks. Barring Question 5 & Question 1(b) (involving Cytoscape),* ***do not*** *use existing network analysis libraries (such as NetworkX/igraph). All results must be submitted through a well-documented Jupyter Notebook.* ***Please submit the data, Jupyter Notebook, the PDF of the final Jupyter Notebook, and (for Question 6 and Question 7) PDF of the scans of the handwritten answers for ease of assessment.***
- Please start working on the assignment right away, as the tasks may take time to complete.

1. Using data of any network of your choice ($N > 200$), **write a Python script** to: (a) Represent the network in terms of its 'adjacency matrix' as well as 'edge list'. (b) Visualize the network. (c) Comment on the 'sparseness' of the network. (d) Compute its average degree $< k >$. (e) Plot its 'scaled degree distribution', $p_k \times k$. (f) Compute its Average Path Length (*Implement Breadth First Search Algorithm*), Diameter and Average Clustering Coefficient. **[15]**

2. **Write Python script** for computing in/out-degree for directed graphs. For a real-world directed network, **compute and plot** its in- and out- degree distribution. **[2]**

3. How would redefine the notion of 'degree' and 'clustering coefficient' for a weighted network to account for the edge weights? **Implement a Python script** to compute these and, for any relevant real-world graph, plot (a) 'weighted degree distribution' and (b) 'Clustering Coefficient' versus 'Degree'. **[5]**

4. **Write a Python script** to create a Gilbert random graph corresponding to an undirected and unweighted real-world network. **Plot and compare** their 'degree distributions'. Compute the degree distribution of the random graph over 100 instances. **[5]**

5. Load the real-world networks studied in the above examples in *Cytoscape* and visualize them using various layouts. Export the images. **[3]**

6. [**Handwritten**] For an undirected and unweighted graph, show that $< k > = {}^{2L}/_{N}$. **[2]**

7. [**Handwritten**] Let A be the $N \times N$ adjacency matrix of an undirected unweighted network, without self-loops. Let 1 be a column vector of N elements, all equal to 1. In other words 1 = (1, 1, ..., 1)^T, where the superscript T indicates the *transpose* operation. Use the matrix formalism (multiplicative constants, multiplication row by column, matrix operations like transpose and trace, etc., but avoid the sum symbol Σ) to write expressions for: **[5×1=5]**

   a. The vector k whose elements are the degrees k_i of all nodes i = 1, 2,..., N.
   b. The total number of links, L, in the network.
   c. The number of triangles T present in the network, where a triangle means three nodes, each connected by links to the other two (Hint: you can use the trace of a matrix).
   d. The vector k_{nn} whose element i is the sum of the degrees of node i's neighbors.
   e. The vector k_{nnn} whose element i is the sum of the degrees of node i's second neighbors.