

Network Science

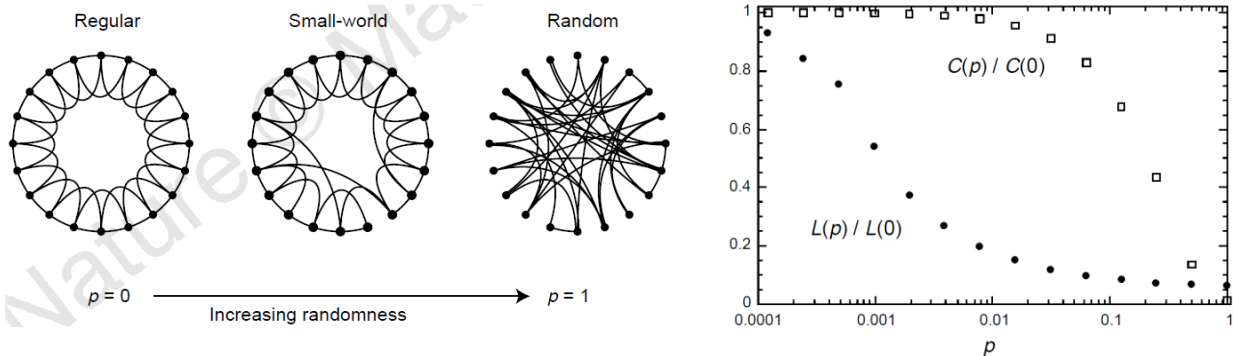
Assignment 2

- All results must be submitted through a well-documented Jupyter Notebook describing the solution to the question in thorough detail.
- You **must** provide the data, Jupyter Notebook as well as the PDF of the final Jupyter Notebook FOR EACH QUESTION SEPARATELY.
- **Delayed submissions will be flatly penalized with 3 points (out of total 10 scaled points).**
- Please start working on the assignment right away as the tasks are computationally intensive.
- The single ZIPped file to be uploaded should be named with following convention: **TeamNumber.zip**

1. Implement Watts and Strogatz's small-world network model. **Compute and plot** the 'scaled clustering coefficient' and 'scaled characteristic path length' of Watts and Strogatz network models with increasing value of rewiring probability. Choose the values of n, k suitably large.

Exact replication of the below-shown plot is expected.

[20]

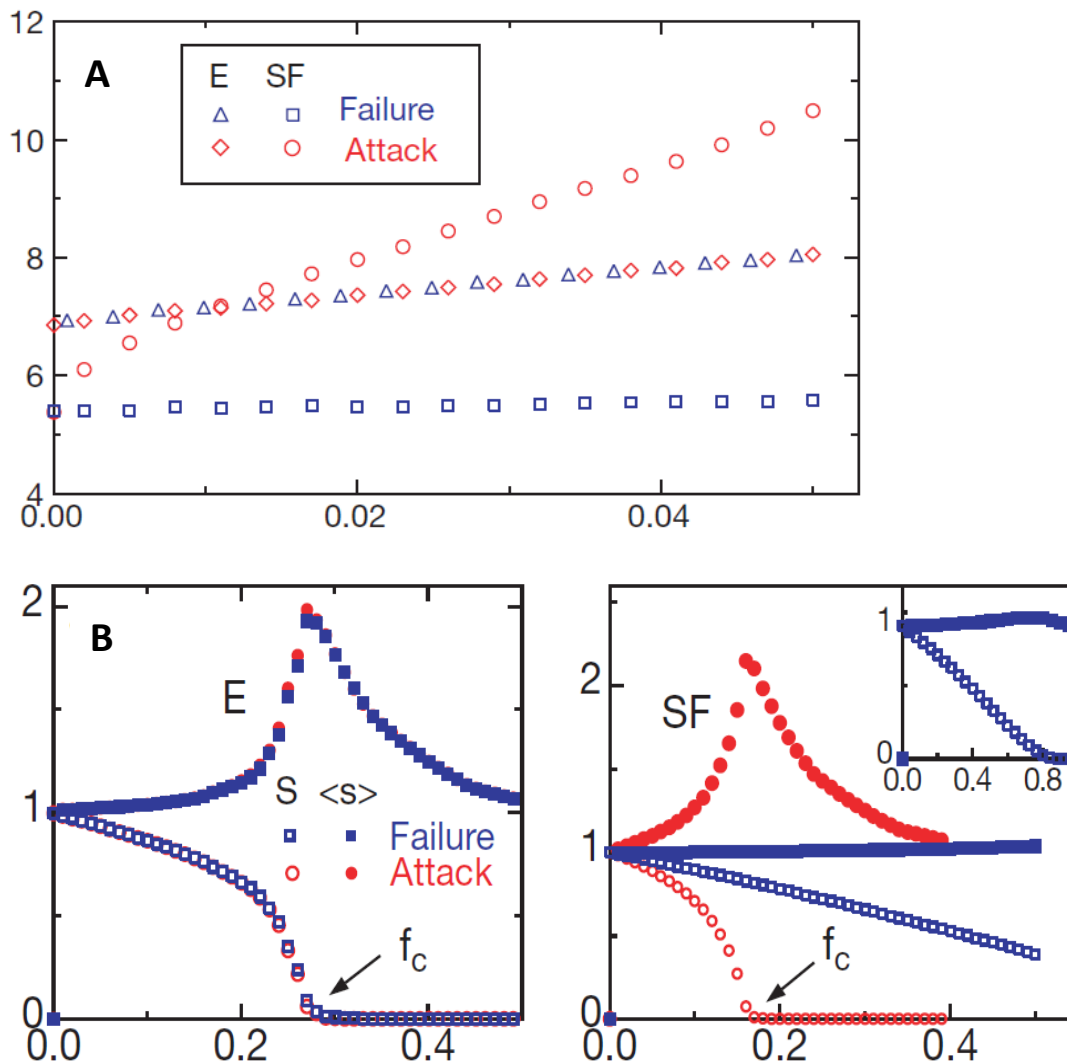


2. (a) **Write the code** to implement Barabasi-Albert (BA) algorithm for generation of scale free networks. Vary the size of the initial random network as well as number of nodes and edges added at every stage of evolution. **Assess the topology** of the final network (minimum 100 instances) in terms of its (a) average clustering coefficient, (b) characteristic path length, and (c) degree distribution. [10]

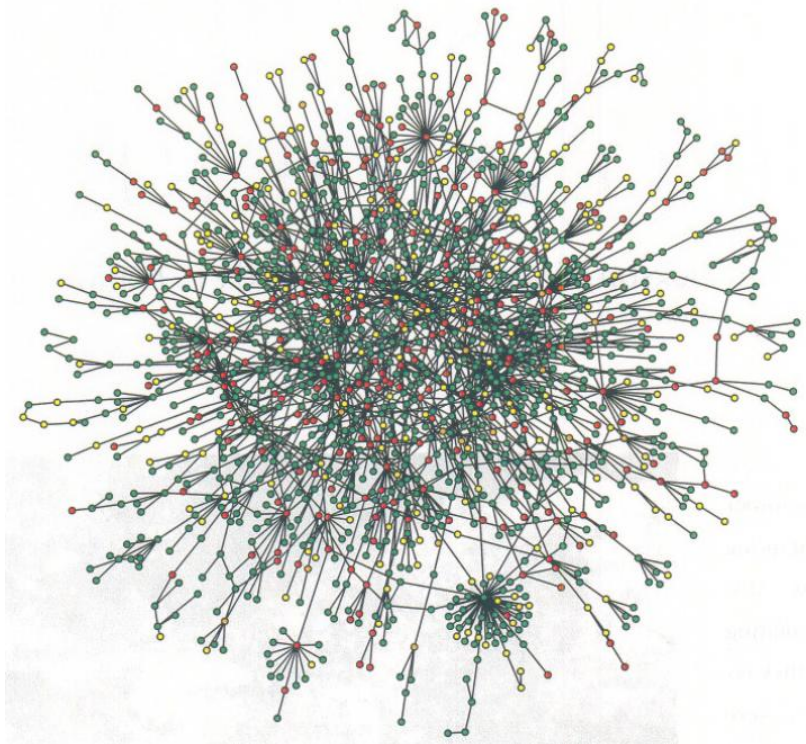
$$P_i(k) \propto \frac{k_i}{\sum_{j=1}^n k_j}$$

3. Modify the Barabasi-Albert algorithm to accentuate/strengthen the bias of rich getting richer phenomenon such that the probability of a newly added node getting connected to an existing node is now "proportional to the square of its degree". **Compute and compare topological features** a comparable size of networks created using BA algorithm. Create variants of higher order. [10]

4. (a) **Implement** the ‘random’ node deletion strategy in a network. Measure the response in terms of ‘characteristic path length’ of the network when increasing fraction of nodes (f) are removed. Also measure the size of the giant cluster (S). You may use existing graph analysis libraries (such as igraph/NetworkX). [5]
- (b) **Extend the study** to compare the response of random and scale free networks to node deletions according to (i) random, and (ii) targeted strategies. [5]
- (c) **Perform** the above-mentioned **analysis on any real-world (scale-free) network** of your choice ($500 < n < 5000$). [5]
- (d) **Comment on your observations**, knowing that the observed pattern of response in (i) characteristic path length and (ii) size of giant cluster/average size of fragmented component is expected to be as follows, as reported by Albert *et al.*, Nature, 406, 378 (2000). [5]



5. (a) Construct the Yeast Protein Interactome (YPI) using data available from Yeast Interactome Database (http://interactome.dfci.harvard.edu/S_cerevisiae/). Plot its degree distribution as well as $C(k) \times k$ plot. Visualize the network using Cytoscape and highlight importance of a protein in terms of its degree/connectivity. [5]
- (b) Identify key proteins in YPI based on their number of interactions. Compute and plot the fraction of proteins that are essential for a given degree (k). (Find the information of essential genes from existing data sources from the internet.) [5]
- (c) Implement the random and targeted node deletion strategies (Refer to Question 3 above) for YPI. Compute and plot the robustness (S and $\langle s \rangle$) of YPI for these strategies. [5]



(‘Lethality and centrality in protein networks’, H Jeong, SP Meson and A-L Barabasi, *Nature* 411, 41-42, 2001.)