# Network Science

# PHYS 5116, Fall 2016

Prof. Albert-László Barabási, Dr. Emma K. Towlson and Dr. Sean P. Cornelius

## Assignment 1, due by October 7th, no later than 6 pm

*Write your name at the top of your assignment before handing it in. Staple all pages together. If you hand in a digital/scanned copy, please be nice to your instructors and hand in a **single**, combined file.*

1. **Matrix formalism**

   Let $\mathbf{A}$ be the $N \times N$ adjacency matrix of an undirected unweighted network, without self-loops, of size $N$. Let be $\mathbf{1}$ a column vector of $N$ elements all equal to 1, that is $\mathbf{1} = (1, 1, \ldots, 1)^T$, where the superscript $T$ indicates the operation *transpose*. In terms of these quantities and by using the matrix formalism (multiplicative constants, multiplication row by column, simple matrix operations like transpose and trace, etc. (no sum symbol $\sum$ allowed!), write expressions for or answer each of the following:

   a) the vector $\mathbf{k}$ whose elements are the degrees $k_i$ of the nodes $i = 1, 2, \ldots, N$;

   b) the total number $L$ of links in the network;

   c) the matrix $\mathbf{N}$ whose element $N_{ij}$ is equal to the number of common neighbors of nodes $i$ and $j$;

   d) the number of triangles, $T$, present in the network. A triangle is three vertices, each connected by edges to both of the others (Hint: you can use the trace of a matrix).

   e) How would you determine whether the network is connected only by looking at the adjacency matrix (following one or more of the operations allowed above)?

2. **Erdős-Rényi graph**

   In an Erdős-Rényi graph with $N = 5000$ nodes, the linking probability is $p = 10^{-4}$.

   a) What is the expected total number of links $\langle L \rangle$?

b) In which regime is the network? Why?

c) Provide the linking probability $p_c$ to have a network with the same number of nodes, but at the critical point?

d) Assuming the same linking probability in (a) and (b), provide the number of nodes $N'$ of a network that (with very high probability) has only one component (one example).

e) For the same example network of (d), calculate the average degree $\langle k' \rangle$ and the average distance among two randomly chosen nodes $\langle d \rangle$.

f) Provide the formula for degree distribution $P(k)$ of this network (approximate as continuous).

3. **Our class network**

   You will find attached a copy of our class network in **edge list format**. Use this representation to answer the following questions about the network. Note: each of these calculations should be done *by hand*, no computers allowed! (Show your work).

   (a) Draw the degree distribution. *Label your axes.*

   (b) Calculate local clustering coefficient for nodes: 9, 14, 19, 21, 25.

   (c) How many components does the network have? What are their sizes? Does any satisfy the definition of a "giant component"? *Hint*: Does this question even make sense?

   (d) For the subnetwork corresponding to the largest component in (c), use the BFS algorithm to make a table containing the shortest path lengths between all pairs of nodes. What is the diameter of this compnent?

4. **Introverts and Extroverts**

   Consider a group of $N$ introverts, who prefer to be alone and hence form an Erdös Rényi network with average degree $\langle k_{\text{int}} \rangle < 1$. Now suppose a small fraction $f$ of the people go to band camp one year and realize they now adore the company of other people, becoming extroverts with degree $k_{\text{ext}} \gg 1$. These extroverts assign their new links randomly, both among themselves and their former co-introverts. You can assume that they also engaged in a form of brainwashing at this particular band camp, erasing any old links of the newly-minted extroverts. You can also assume that as a result of going to camp together, there is a "free" link between any pair of extroverts.

(a) What is the distribution of connected component sizes (in terms of $N$ and $k_{\text{int}}$) for the original network of only introverts? Note: as a distribution, your function should be properly normalized!

(b) What is the minimal extrovert fraction $f$ required to ensure (with high probability) the network is connected? You should express your answer only in terms of $N$, $k_{\text{ext}}$ and $k_{\text{int}}$. *Hint:* Think about part (a). What is the minimum the extroverts have to do to make the network globally connected? Explain and interpret your results. What happens as $k_{\text{in}}$ becomes very small? What happens as $k_{\text{in}}$ approaches 1?

(c) For large networks satisfying the criteria derived in (b), what is the expected shortest path length between two introverts? An introvert and an extrovert? Do these networks have the small world property?

5. **Bipartite networks**

Problem 2.12.5 in the textbook, plus part (d) of Problem 2.12.6.