# Network Science
# PHYS 5116, Fall 2015

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

## Assignment 3, due by November 28th, 5.30pm (*i.e.*, before class)

**You are only allowed to hand in a single file containing all your work, either a PDF or a stapled paper copy. If you attempt to hand in more than one file, we will only grade one selected at random.**

1. **Conspiracy in Social Networks**
In a Big Brother society, the thought police wants to follow a "divide and conquer" strategy by fragmenting the social network into isolated components. You belong to the resistance and want to foil their plans. There are rumors that the police want to detain individuals that have many friends as well as individuals whose friends tend to know each other. To decide which people to protect, you simulate two different attacks on your network by removing, iteratively, (i) the node with the highest clustering coefficient and (ii) the node with the largest degree. For each type of attack, study (by computer simulation) the size of the largest component as a function of the fraction nodes removed for each of the following networks:

    a) A random network with $10^4$ nodes generated with the Barabási-Albert model with parameter $m = 10$. You may use an existing library function (such as `barabasi_albert_graph` in NetworkX) to generate the network, or your implementation from the last homework (provided it was correct!).

    b) A random network generated the hierarchical model described in Figure 9.13and ADVANCED TOPIC 9.A, using 6 iterations of the algorithm. You will have to implement this algorithm yourself. *Hints:* You can use `complete_graph` function in NetworkX to generate cliques of size 5. You should also make judicious use of the `disjoint_union` and `copy` functions for graphs. You can easily keep track of whether a node is a "hub" vs. "peripheral" (and hence whether you should connect it to the central "big brother" node) by using a node attribute (*e.g.* `G.node[u]['peripheral'] = True`. This attribute will be preserved when you make 4 `copy`-s of a graph...

    For each graph, which is the most sensitive piece of topological information: clustering coefficient or degree? In other words, which one–if protected–limits the damage best? Would it be better if all individuals' information (clustering coefficient, degree, etc.) could be kept completely secret? Why? You should include your code and four plots—one for each attack type/network pair—with your solution.

2. **Evolving networks** Consider the model presented in problem 6.7.2 in the book. Assume that the fitnesses are distributed with probability $p(\eta)$. Approach this problem in the following way:

    (a) Using the continuum theory, what is the rate equation $\left(\frac{dk_i}{dt} =?\right)$ for the time evolution of the degree $k_i$ of node $i$?

    (b) Integrate both sides of this equation with respect to $t$ (between the time the node was born, $t_i$, and the time now, $t$) to obtain $k_i(t)$, the degree of node $i$ as a function of time $t$.

    (c) Following the approach for the BA model (Box 5.3), calculate the number of nodes with fitnes $\eta$ having degree less than or equal to a given value, $k$. Your answer should be a function only of $\eta$, $N$ and $k$.

    (d) Divide this by the total number of nodes of fitness $\eta$, namely $N_\eta = Np(\eta)$, to obtain the cumulative distribution function (CDF) of degree for the subset of nodes with fitness $\eta$. Take a derivative with respect to $k$ to obtain the associated probability density function (PDF).

    (e) Now, assuming that the fitness distribution is half nodes with $\eta = 1$, and the other half $\eta = 2$, what is the *overall* degree distribution of the network? Relate the result of this model to the ER model. What is the average degree of a "good-looking" node? A less-good-looking node?

3. **Motifs**

a) How many undirected motifs of size 4 exist? Draw them. How many directed motifs exist? (Don't draw them!)

b) Consider a random network created via the $G(N, p)$ model, with $N = 10000$ and $p = 0.004$. For each of the undirected 4 node motifs you identified in part (a), calculate the expected number of occurrences of the motif in this network.

4. **Bianconi-Barabási model (compulsory for Network Science PhD students only)**
Consider a network growing with the Bianconi-Barabási model where nodes have two distinct fitnesses, $\eta = a$ and $\eta = 1$. To be specific, let us assume that the fitness follows the double delta distribution

$$\rho(\eta) = \frac{1}{2}\delta(\eta - a) + \frac{1}{2}\delta(\eta - 1) \text{ with } 0 \leq a \leq 1 \tag{1}$$

a) Calculate the degree exponent, and how it depends on the parameter $a$.

b) Calculate the stationary degree distribution of the network.