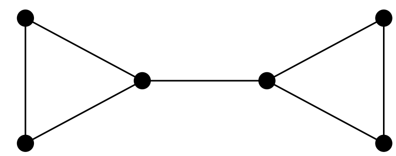
Complex Systems 535/Physics 508: Homework 8

Because of the Thanksgiving break you have longer than usual to do this homework. It is due in class on ***December 5****.*

1. Using your favorite numerical software for matrices, construct the modularity matrix for this small network:



Find the eigenvector of the modularity matrix corresponding to the largest eigenvalue and hence divide the network into two communities. For full credit, give the eigenvector you found and draw a figure showing how it divides the network into communities.

2. Consider the following network model. Each of n nodes is assigned a non-negative real parameter θ_i and then undirected edges are created such that the number of edges between nodes i and j is a Poisson-distributed independent random number with mean $\theta_i\theta_j$, except for self-edges, which have mean $\frac{1}{2}\theta_i^2$. The goal is to find the values of the parameters θ_i that best fit a given observed network.

   (a) Derive an expression for the likelihood (i.e., the probability) that a network with adjacency matrix $\mathbf{A}$ was generated by this model, given the values of the parameters θ_i, and hence show that the log-likelihood $\mathscr{L} = \log P(\mathbf{A}|\boldsymbol{\theta})$, ignoring constants that don't depend on θ_i, is

$$\mathscr{L} = \frac{1}{2}\sum_{ij}\big[A_{ij}\log(\theta_i\theta_j) - \theta_i\theta_j\big].$$

   (b) By maximizing the log-likelihood with respect to the θ parameters show that for the best fit of this model to an observed network, the mean number of edges between distinct nodes i and j is $k_ik_j/2m$, where k_i is the observed degree of node i and m is the number of edges in the network. (In other words, this model is basically the configuration model.)

3. Consider a configuration model network that has nodes of degree 1, 2, and 3 only, in fractions p_1, p_2, and p_3, respectively.

   (a) Find the value of the critical node occupation probability ϕ_c at which site percolation takes place on the network.

   (b) Show that there is no giant cluster for any value of the occupation probability ϕ if $p_1 > 3p_3$. In terms of the structure of the network, why is this? And why does the result not depend on p_2?

   (c) Find an expression for the size of the giant cluster as a function of ϕ. (Hint: you may find it useful to remember that $u = 1$ is always a solution of the equation $u = 1 - \phi + \phi g_1(u)$.)

4. Consider the problem of nonuniform percolation on a configuration model network, as discussed in Section 15.3 of the course pack, where the occupation probability ϕ_k for a node is a function of degree k.

   (a) Show, by a graphical argument or otherwise, that a giant component can exist in the network only if $f_1'(1) > 1$, where $f_1(z)$ is the function defined in Eq. (15.34).

   (b) A configuration model network has a (properly normalized) geometric degree distribution $p_k = (1-a)\,a^k$ with $a < 1$ and an occupation probability $\phi_k = b^k$ with $b < 1$, so that high-degree nodes are more likely to be removed than low-degree ones. Show that the system has a giant cluster if $2ab^2(1-a)^2 > (1-ab)^3$.

5. **Extra credit:** Write a program in the computer language of your choice to do the following:

   (a) Generate a Poisson random graph of the type $G(n, p)$ with $n = 10\,000$ nodes and mean degree $c = 4$.

   (b) Carry out the percolation algorithm of Section 15.5 on this network, occupying the nodes one by one and updating cluster labels after each one. Keep a running tally of the size of the largest cluster.

   (c) Repeat the entire calculation 100 times for 100 different random graphs with the same value of c and produce a plot showing the mean (over the 100 runs) of the size of the largest cluster as a function of the number of occupied nodes.

   (d) Based on this plot, at about what value would you say the percolation threshold falls?

   For full extra credit, turn in a printout of your complete program, a printout of the plot you produced, and your answer to the question in part (d).