Math 168, Networks, UCLA, Spring 2018
Problem Sheet 7: The Final Countdown
(i.e., the final homework assignment)

(submit to CCLE by Wednesday 30 May 2018 at 5:00 pm)

1. *Reading.*
   (a) Read Chapters 10, 11, and 16 of Newman's book.
   (b) Read the article *Communities in Networks* (by Porter, Onnela, and Mucha): `http://www.math.ucla.edu/~mason/papers/comnotices.pdf`
   (c) Read the booklet "Dynamical Systems on Networks: A Tutorial" (2016) by Porter and Gleeson. There is a copy of it in CCLE.
   (d) Note: For more details on community structure, you may also wish to browse through the article *Community Detection: A User Guide* (2016) by Fortunato and Hric and *Community Detection in Graphs* (2010) by Fortunato. Those of you who are doing clustering as part of your project will want to look at the Fortunato–Hric article.
2. Do Problem 12.3 of Newman's book.
3. Do Problem 13.1 of Newman's book.
4. Do Problem 13.4 of Newman's book.
5. *Generalization of configuration model and hypergraphs.*
   (a) Define *hypergraph*, and indicate how a hypergraph can be represented as a bipartite network.
   (b) Consider a generalization $\mathcal{G}$ of the configuration model in which there are three different types of nodes (red, green, and blue) and in which each adjacency among a set of nodes is determined by a hyperedge that is incident to exactly one node of each of the three types.
      (i) Write an explicit formulation for this random-graph ensemble.
      (ii) Indicate one example of a real-world network that such a hypergraph can be used to model. State explicitly what each type of node is.
      (iii) Suppose that $g_0(z) = \sum_{k=0}^{\infty} p_g(k) z^k$ is the generating function for the degree distribution of green nodes, and let $g_1(z)$ denote the generating function for the corresponding excess degree distribution. Derive an expression for $g_1(z)$ in terms of g_0 and its derivatives.
      (iv) Consider the projection of a network from the random-graph ensemble $\mathcal{G}$ onto red nodes, such that two red nodes are adjacent in this projection if they share a green neighbor. Let $\rho_g(k)$ denote the probability that a red node a has k neighbors in the projected network. Show that the corresponding generating function $R_g(z)$ satisfies the equation $R_g(z) = r_0[g_1(z)]$, where $r_0(z)$ is the generating function for the degree distribution of red nodes.
      (v) Now consider a projection in which a pair of red nodes are adjacent if they share either a green neighbor or a blue neighbor. Derive an expression for the mean degree in this projection in terms of generating functions for the degree distributions and excess degree distributions of the red, green, and blue nodes.
6. *Community structure in networks.*
   a. Consider two networks from empirical data. (Make sure they are of sufficiently different sizes, so perhaps one has about 5–10 times as many nodes as the other.) For each of these two networks, detect communities algorithmically in at least two different ways. One of these ways should be a form of modularity maximization, and you can choose any other method for the second one. Compare the results of these methods to each other.
   b. Consider a "ring-of-pearls" network that is unweighted and undirected. It consists of some number α of k-cliques, such that there is exactly one edge between each pair of k-cliques. Fix the value of k to be a reasonably large number. For different values of α, what do you get as a result of modularity maximization at the default resolution-parameter value. What happens as you adjust the resolution parameter? What happens if you generalize this network to consider "pearls" with heterogeneous values of k? What do your results imply?

c. Perform community detection for a collection of Erdős–Rényi random graphs. What is the result, and what does it imply?