Assignment 5 (Graph Theory and Networks)
Discussed in class on April 16

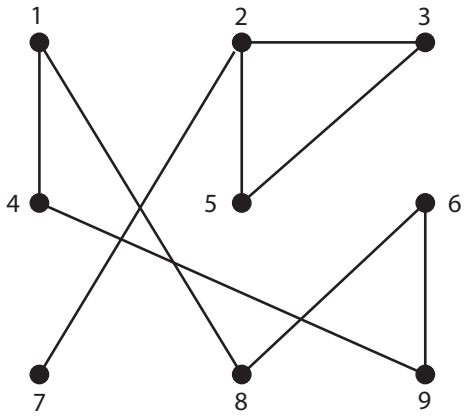
---

(1) The graph below is missing one edge. With the edge present, the graph has the following clustering coefficients: $C_1 = \frac{1}{3}$, $C_2 = \frac{1}{3}$, $C_3 = 1$, $C_4 = 0$, $C_5 = 1$, $C_6 = \frac{1}{3}$, $C_7 = 0$, $C_8 = 1$, $C_9 = 0$. (Note that a node with only one neighbor has clustering coefficient of 0.)

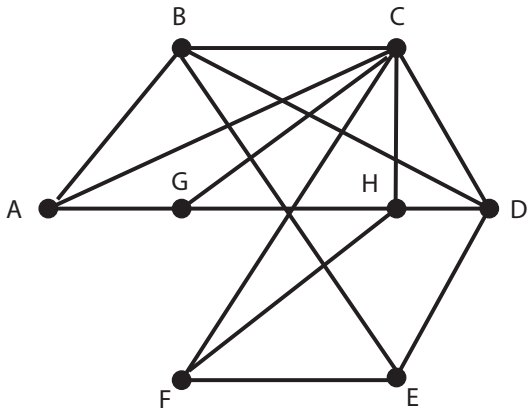
(a) Which nodes should be connected by an additional edge to make these clustering coefficients correct?

(b) Which nodes should be connected by an additional edge so that $C_7 = 1$? What impact would this have on the other clustering coefficients if it is added?

(c) Erase all the edges and connect the nodes with edges so that each node has degree of at least 2 and clustering coefficient of 0.



(2) In the social network below, calculate the neighborhood overlap of all edges and place tie strengths on the edges in such a way that greater tie strengths are associated with greater neighborhood overlap. Are there any local bridges? If so, list them.



(3) Consider a ring-like lattice like the one used in the Watts-Strogatz model, with no random rewiring. Each node connects with K nearest neighbors; half on one side and half on the other. Determine the mean clustering coefficient for the three cases $K = 2, 4, 6$ and in each case explain your derivation.

(4) Consider an Erdős-Rényi random graph G(10,0.5).
   (a) What is the mean number of edges $< m >$?
   (b) What is the mean degree c?
   (c) What is the probability that a node has degree 0?
   (d) What is the probability that a node has degree 9?

(5) For an Erdős-Rényi network the maximum geodesic path length (also called the *network diameter*) is approximately $\frac{\ln n}{\ln c}$ where c is the mean degree. If such a network has 1000 nodes, then what must p be so that the diameter is approximately equal to 2?

(6) Consider two versions of the rank model, with different ranking criteria. In the first version nodes are ranked by age (the time elapsed since they were added to the network). In the second version nodes are ranked by their degree. Is there a difference between networks generated by the two models, and if so, what is it?