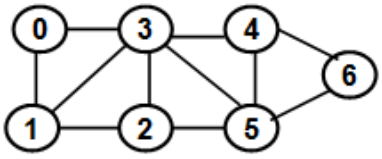
CSC 641 Network Science, Fall 2015
Instructor: Dr. Natarajan Meghanathan

Take Home Exam 2
Due: Monday, Nov. 2, 2015: 6 PM

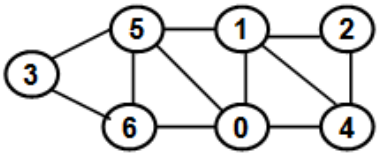
Late submission (Nov. 2, 2015-6:10 PM to Nov. 3, 2015-6 PM: -30 points, taken off from your score)
Late submission (Nov. 3, 2015-6:01 PM to Nov. 4, 2015-6 PM: -60 points, taken off from your score)
No late submission allowed after Nov. 4, 2015-6 PM.

Maximum Points: 100

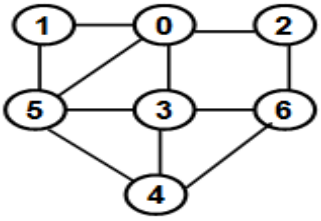
1) (a-16 pts) Run the "complete linkage clustering" algorithm to determine a hierarchy of clusters
(b-12 pts) Compute the modularity scores of the communities at each level of the hierarchy and determine the partition of vertices into communities with the largest possible modularity score
(c-12 pts) Compute the Silhouette Index of the vertices to the communities with the largest modularity score identified in (b) and conclude whether each vertex is in the appropriate community or not.



Karthik Reddy

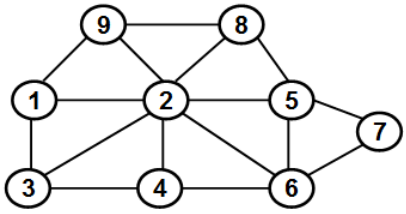


Anirudh Reddy

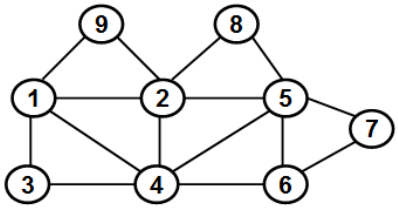


Yashvanth Divanji

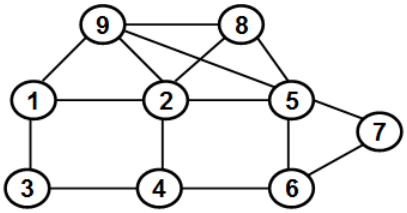
2) (12 pts) Consider the graph assigned to you.
(a) Determine the embeddedness of each edge and vertex of the graph.
(b) Determine the structural hole(s) of the network - the nodes with the smallest embeddedness. Does the graph get disconnected into two or more community if the structural holes and the associated edges are removed?



Karthik Reddy



Anirudh Reddy

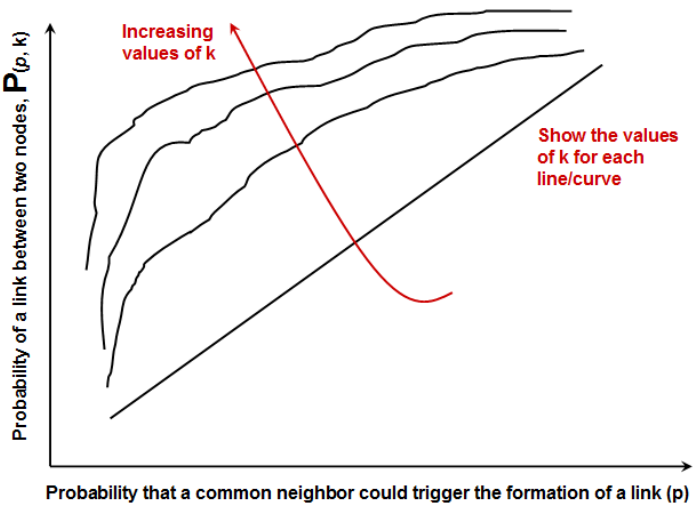


Yashvanth Divanji

3) (16 pts) Run the Eigenvector-based community detection algorithm on the graphs assigned to you in Q4 above and determine the partition of vertices into communities with the largest possible modularity score.

See next page for questions 4 and 5.

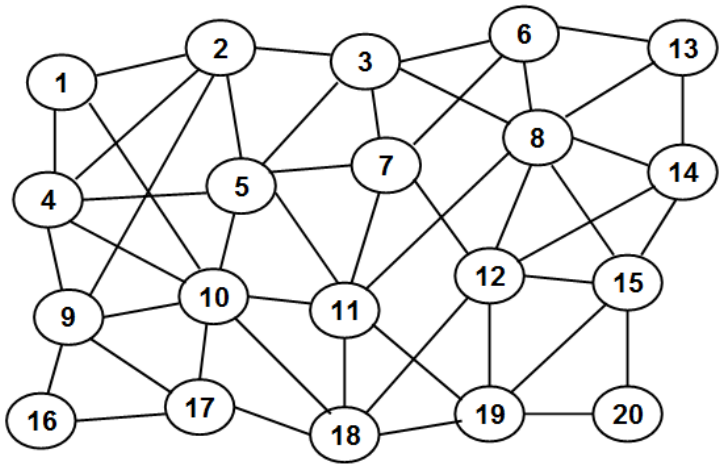
4) (14 pts) Let there be two nodes that are initially not connected. There are ***k*** common neighbors for the two nodes. Let ***p*** be the probability that the presence of a common neighbor could trigger the formation of a link. Plot an Excel graph that shows the probability (**P**) that the two nodes will be connected as a function of ***p*** and ***k***. Vary the values of ***p*** from 0.1 to 1, in increments of 0.1 and vary the values of ***k*** from 1 to 7, increments of 1. You should show the plot in one single figure like this.



Also, for each of you, I have assigned a value of **P**(***p***, ***k***) and ***p***. Given the value of ***p***, you will have to find the minimum value of ***k*** that will contribute to a probability of at least **P**(***p***, ***k***) for a link to exist between the two nodes. Use your Excel data and plot.

Student	Probability that a common neighbor could trigger the formation of a link (***p***)	Probability that a link exists between the two nodes is at least **P**(***p***, ***k***)
Karthik Reddy	0.4	0.75
Anirudh Reddy	0.3	0.70
Yashvanth Divanji	0.2	0.65

5) (18 pts) Assume the graph given below comprises of three sets of non-overlapping vertices representing people of Asian, Caucasian and Hispanic origin within a neighborhood. The edges represent friendships between these people.



(a) Count the total the number of edges in the graph based on the node degree. Show your work. Do not count the total number of edges manually. [Review the beginning portion of Module 1 slides, if needed]

Determine the following using the Homophily-approach:
(b) Whether the Asians and Hispanics exist as two separate communities or as one single community?
(c) Whether the Asian and Caucasians exist as two separate communities or as one single community?
(d) Whether the Hispanics and Caucasians exist as two separate communities or as one single community?

Each of you are assigned a distinct grouping of the vertices into the above three sets. Your calculations have to be based on this grouping.
Also, I have posted a powerpoint version of this graph in the course website. You could use the powerpoint slide to color the vertices and do your calculations.

Student	Asian origin vertices	Caucasian origin vertices	Hispanic origin vertices
Karthik Reddy	5, 7, 8, 10, 11, 12	1, 2, 3, 4, 6, 13, 14	9, 15, 16, 17, 18, 19, 20
Anirudh Reddy	1, 2, 3, 4, 5, 7	9, 10, 11, 16, 17, 18	6, 8, 12, 13, 14, 15, 19, 20
Yashvanth Divanji	3, 6, 7, 8, 13, 14	11, 12, 15, 18, 19, 20	1, 2, 4, 5, 9, 10, 16, 17