Question: 4) Is there a strong correlation between people connected by the adjacency graph and people connected by interest in common research topics? What is a measure you would propose to quantify this?

Solution: Here is a list of measures people devised:

i) Consider the bipartite graph of people-to-topics built up in problem 3). Now add people-to-people links between any two individual people who are acquainted. (This makes a hybrid of both graphs considered thus far). Now compute the clustering coefficient *for the topics* (i.e., if people X and Y are both interested in topic A, how likely is it that X and Y are linked in the acquaintance network?).

ii) Consider the "correlation coefficient" for each node (the number of friends with the same interest, divided by the total number of friends for that node). Start with an individual X, and their degree, d_X. Now consider the number of friends of X who have the same research interest as X, call this r_X. A measure for the overlap for X is r_X/d_X. Now take the average value of r_X/d_X over all people in the network.

iii) Now consider a complementary measure — the number of people in a given research component who are friends, divided by the total number of people in that research component. (Note above we considered the number of friends with the same interest, divided by the total number of friends for a node). Compute the average of this measure over the whole network.

iv) Consider the two graphs 1) the people-to-people (social) graph (call that graph S), and 2) the people-to-research_topics graph (call that graph R). Look at the shortest path connecting people in S, and the shortest path in R and calculate the correlation coefficient between them.