NAME	NIA	GRADE

Introduction to Network Science (2018-2019)

FINAL EXAM

WRITE YOUR ANSWERS CLEARLY IN THE BLANK SPACES. Write as if you were trying to communicate something in written to another person who is going to evaluate what you write. If for some reason (for example, if after you have written the solution you realize that there is some mistake that you would like to correct) you can attach an extra sheet to your exam. In this case, indicate clearly that the solution can be found in the extra sheet. Also, you may use other sheets to perform your calculations.

Problem 1 *1 point*

Suppose we want to design vaccination strategies to prevent a disease from reaching a large number of people.

1. Indicate how to choose whom to vaccinate if we do not know fully the social network, and want to base the vaccination strategy on the friendship paradox we saw in class.

2. Indicate how to choose whom to vaccinate if we know fully the social network and want to base the vaccination strategy on the Independent Cascade model.

Problem 2 *1 point*

Describe step by step in pseudocode how to create a Barabási-Albert graph with N nodes having m_0 starting nodes and m outlinks per node. For your pseudocode to be valid, if at any point there is a randomized step, you must indicate what is the probability of each possible outcome.

Problem 3 *1 point*

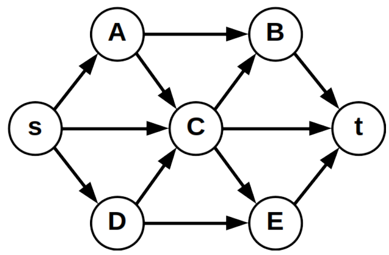
Under a preferential attachment model with aging effects, the probability of a node with degree k_i created at time t_i to gain a new link at step t is $\Pi(k_i, t - t_i) \approx k_i(t - t_i)^{-v}$ where v is an exponent controlling the aging process.

1. What is the value of v for this to be equivalent to standard preferential attachment?

2. When $v \to \infty$ we observe that the resulting graph is a line, this is, a sequence of nodes in which each one is connected only to the ones preceding and following it. Why?

Problem 4 *1 point*

Consider $(s, t)-$cuts on the graph on the right, where s is the source node and t is the terminal node. Assume every edge has cost equal to 1.

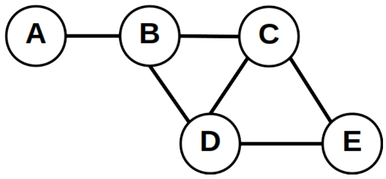


1. By visual inspection, what is the minimum cost of an $(s, t)-$cut in this graph, and what is an example of a cut having that cost? (Feel free to draw it over the graph on the right)

2. Run the algorithm for randomized (s, t)-cuts we saw in class, drawing all intermediate graphs, and indicate the cost of the resulting cut.

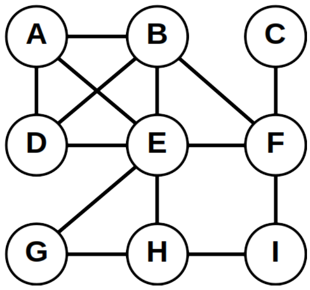
Problem 5 *2 points*

Run Brandes-Newman algorithm to determine edge betweenness on the graph shown on the right. Indicate the final edge betweenness values clearly.



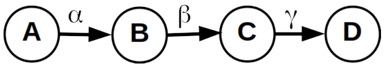
Problem 6 *1 point*

Consider the graph on the right. Run Charikar's randomized algorithm for densest subgraph, indicating all intermediate graphs and their density, and marking clearly the graph with the largest density. For density use $|E|/|V|$ where $|E|$ is the number of edges in the subgraph and $|V|$ the number of nodes in the subgraph.



Problem 7 *2 points*

Consider the graph on the right, including the infection probabilities indicated in the edges: α, β and γ. Let X_i be the expected number of nodes infected under the Independent Cascade Model for an infection starting at node i, including the node initially infected.



For instance, if an infection starts from node B, the probability that the number of nodes infected is 2 is $P(X_B = 2) = \beta \cdot (1 - \gamma)$. This is because for the infected to be 2 we need the infection from B to C to succeed and the infection from C to D to fail.

Remember that the expectation of a variable X is $E[X] = \sum x \cdot P(X = x)$, where the summation is done over the possible values x that the variable can take.

1. What is $E[X_C]$ as a function of γ?

2. What is $E[X_A]$ as a function of α, β, γ?

Problem 8 *1 point*

Under the SIS model (Susceptible - Infected - Susceptible), the proportion of infected nodes $i(t)$ follows

$$i(t) = \left(1 - \frac{\mu}{\beta \langle k \rangle}\right) \frac{Ce^{(\beta\langle k \rangle - \mu)t}}{1 + Ce^{(\beta\langle k \rangle - \mu)t}} ,$$

where μ is the recovery rate, β the infection probability, $\langle k \rangle$ the average degree of nodes, and C a constant that depends on i_0, the initial number of infected nodes.

In the limit when $t \to \infty$:

1. When $\mu < \beta \langle k \rangle$, what is the limit of $i(t)$?

2. How is this state called?

3. What happens when $\mu > \beta \langle k \rangle$?

4. Under which conditions in a SIS model the number of infected nodes is larger?