NAME	NIA	GRADE

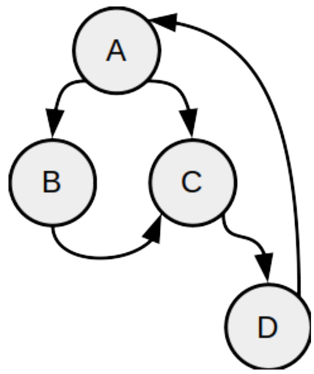
Introduction to Network Science (2021-2022)

——————— *FINAL EXAM* ———————

WRITE YOUR ANSWERS CLEARLY IN THE BLANK SPACES. 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. Please underline or highlight the important keywords of your answers and the final answers when you have performed a calculation. Simply fractions in your answers, if answering with fractions.

Problem 1 *1 point*

Perform two iterations of simplified PageRank for the graph shown on the right. Indicate in the table below the initialization scores, the scores after the first iteration, and the scores after the second iteration.



Answer:

	Init	Iter 1	Iter 2
A			
B			
C			
D			

Problem 2 *1 point*

In class, we discussed extensively that in a directed weighted graph having a node s having only out-links and a node t having only in-links, the maximum flow that can be carried from s to t is equal to the minimum cut separating s from t.

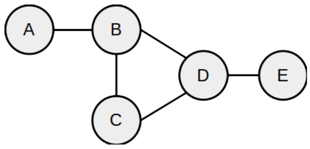
Next, (1) briefly explaining why this equality holds, in your own words, (2) give an example graph with both the min cut indicated as a line crossing some edges, with its cost, and the max flow indicated with numbers representing flows in each edge of the graph.

Brief explanation of why this equality holds, in your own words; you can use formalism but it is not necessary:

Example graph with min cut and max flow indicated:

Problem 3 *2 points*

Execute the Brandes-Newman algorithm to compute the edge betweenness for the graph shown on the right. Highlight clearly your final answer, which should include the edge betweenness of each edge.



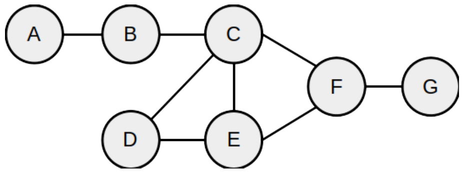
Answer:

Problem 4 *1 point*

Indicate to which k-core the following nodes belong to on the graph on the right:

Answer:

Node	k-core	Node	k-core
A		E	
B		F	
C		G	
D			

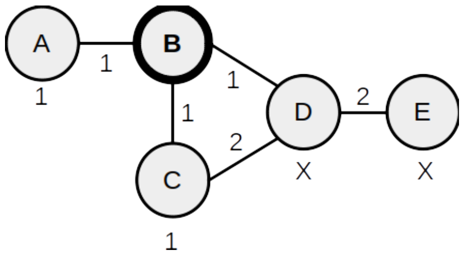


Problem 5 *1.5 points*

For the same graph of the previous problem, run Charikar's randomized algorithm for densest subgraph and indicate the density of the dense subgraph found ($|E|/|V|$, where E is its set of edges and V its set of vertices). Remember to indicate the density of each intermediate graph during the algorithm execution, and highlight the densest one.

Your answer:

Problem 6 *1 point*



Consider the Linear Threshold model of propagation on the figure on the right. Edge weights are written next to each edge, node thresholds are written under each node, and the infection starts from node B.
What is a possible value of X so that both D and E are infected? Explain.

What is a possible value of X so that D is infected but E is not infected? Explain.

Problem 7 *1.5 points*

Suppose in a graph $G = (V, E)$ you want to compute the *influence* of sets of nodes, denoted by $f(S)$ for set $S \subseteq V$, which is the *expected* number of nodes infected under the Indepentent Cascade Model if an infection starts at *all* nodes in S simultaneously.
Indicate how could you compute $f(S)$*:*

Indicate what is the relationship $(<, \leq, =, \geq, >)$ *between* $f(A \cup B)$ *and* $f(A) + f(B)$.

Explain why is that the relationship, and give one example with a small graph, explaining the possible infection process

Problem 8 *1 point*

Consider a model we will call $SIRS$ in which nodes start as being susceptible, then susceptible nodes might get infected, infected nodes may recover, and recovered nodes migt become susceptible again. Let us denote by $s(t), i(t), r(t)$ the fraction of susceptible, infected, and recovered nodes in the graph respectively. At each timestep, with probability β a susceptible node may get infected if in contact with an infected node, with probability μ an infected node might recover, and with probability σ a recovered node might become susceptible again.

Write the differential equation for $s(t)$, explaining each term of the equation

Write the differential equation for $i(t)$, explaining each term of the equation

Write the differential equation for $r(t)$, explaining each term of the equation

Is $\sigma > \mu$ sufficient to say that the recovered will tend to zero in the long run? Explain