Fall Semester 2021

KAIST AI607
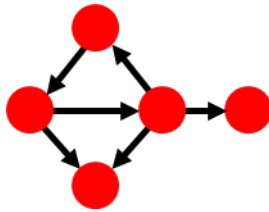
Graph Mining and Social Network Analysis

Homework 3: Take-Home Midterm Exam

Release: October 8, 2020

Due: October 22, 2020, 11:59 PM

1. (10 points) Basic Graph Theory

Consider the following directed graph G:

(a) **(5 points)** How many weakly connected components (WCCs) does G have?

Hint: Recall that WCCs should be maximal. Adding any node to a WCC should make it no longer a WCC.

(b) **(5 points)** How many strongly connected components (SCCs) does G have?

Hint: Recall that SCCs should be maximal. Adding any node to a SCC should make it no longer a SCC.

2. (15 points) Random Graphs

Assume we generate an undirected graph using the random graph model $G(N, p)$ with $N = 4$ and $p = 0.5$

(a) **(10 points)** What is the probability that the generated graph has the following structure? Feel free to write an equation as your answer
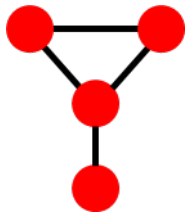

Hint:
- N is the number of nodes
- p is the probability that each pair of two distinct nodes are connected
- There can be multiple graphs with the same structure

(b) **(5 points)** What is the expected value of the average degree of the generated graph?

3. (15 points) Deterministic Kronecker Graph

Assume that an undirected graph G_3 with the following adjacency matrix is generated by the deterministic Kronecker model.
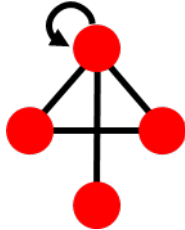

<Sparsity pattern of the adjacency matrix of G_3>
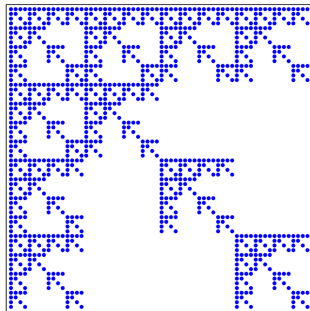
(a) **(10 points)** What is the adjacency matrix of the seed graph?

(b) **(5 points)** What is the diameter of G_5?

4. (15 points) Stochastic Kronecker Graph

Assume the stochastic Kronecker model generates an undirected graph G_2 with the following structure

(a) **(10 points)** If the seed matrix is $\begin{bmatrix} 1 & 0.5 \\ 0.5 & 0 \end{bmatrix}$, what is the probability that the generated graph has the above structure?

(b) **(5 points)** Give a seed matrix that maximizes the probability of generating a graph with the above structure.

5. (15 points) PageRank

Consider the following unweighted directed graph G:
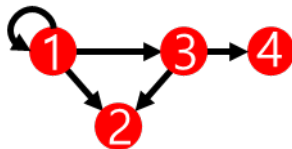
(a) **(5 points)** What is the adjacency matrix A of G?

(b) **(10 points)** Computing $\vec{x} \leftarrow \mathbf{B} \times \vec{x}$ repeatedly makes $\vec{x}$ converge to the PageRank score vector (with damping factor 0.85) of G. What is the matrix $\mathbf{B}$?

    Hint:
- Consider the version of PageRank where the dead-end problem and the spider-trap problem are solved by teleports.
- The probability that the random surfer teleports is (1 - damping factor)
- When the random surfer teleports, it jumps to a uniformly random node.

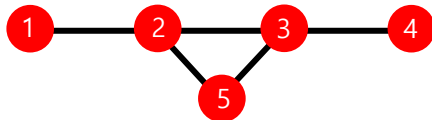## 6. (10 points) Harmonic Centrality / Closeness Centrality

Consider the following undirected graph G:

(a) **(5 points)** What is the harmonic centrality of node 3?

(b) **(5 points)** What is the closeness centrality of node 3?

7. (20 points) Betweenness Centrality / Coreness

Consider the following undirected graph G:

(a) **(5 points)** What is the betweenness centrality of node 5?

(b) **(5 points)** What is the betweenness centrality of node 3?

(c) **(5 points)** What is the coreness of node 5?

(d) **(5 points)** What is the coreness of node 3?

[Notes]

If you encounter any subtleties, ask questions on CLASSUM.

For each question, provide an answer and also explain how you arrived at your answer.

[How to submit your assignment]

(a) Merge your answers into a single pdf file named hw3-[your student id].pdf

(b) Submit the pdf file at KLMS