

Seoul National University

M1522.001400 Introduction to Data Mining

Spring 2016, Kang

Homework 9: Mining Social-Network Graphs (Chapter 10)

Due: June 8, 09:30 AM

Reminders

- The points of this homework add up to 100.
- Like all homeworks, this has to be done individually.
- Lead T.A.: Jinhong Jung (montecast9@gmail.com)
- Please type your answers in English. Illegible handwriting may get no points, at the discretion of the graders.
- If you have a question about assignments, please upload your question in eTL.
- If you want to use slipdays or consider late submission with penalties, please note that you are allowed one week to submit your assignment after the due date.

Remember that:

- Whenever you are making an assumption, please state it clearly

Question 1

For the example graph from Figure 1, use the Girvan-Newman approach to find the fraction of shortest paths from each of the following nodes that pass through each of the edges when we start BFS from the following node. [25 points]

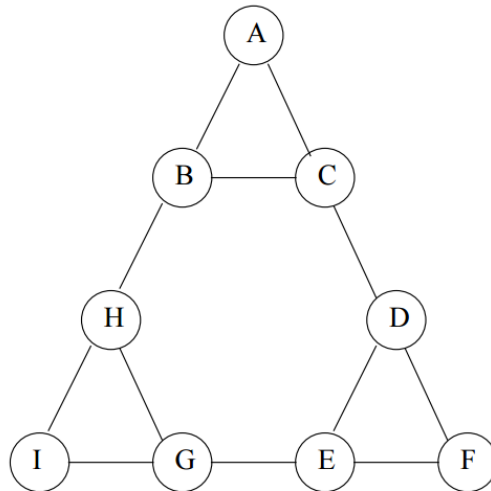
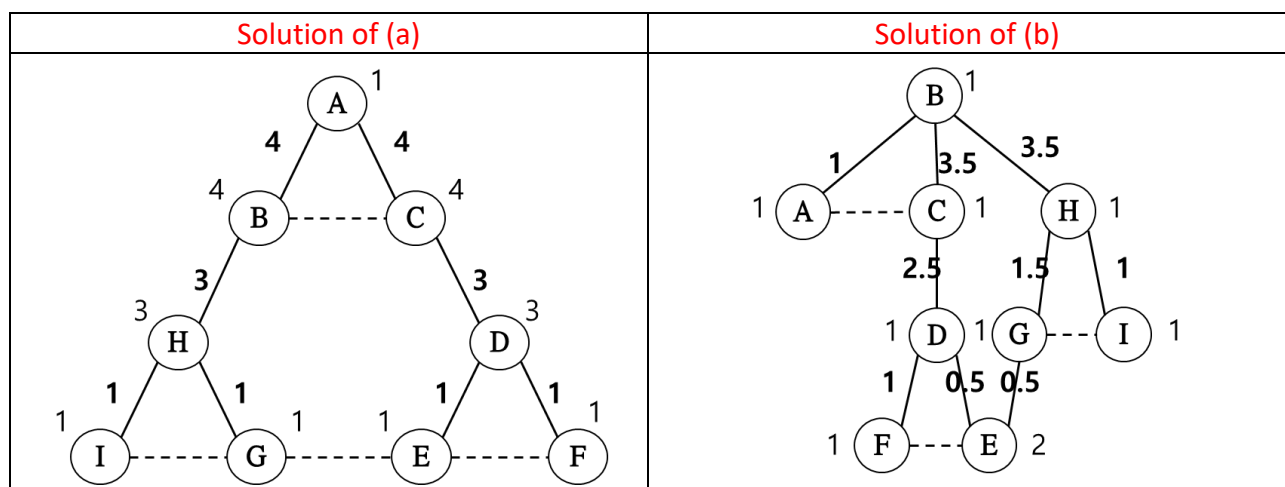


Figure 1. Example graph

(a) A

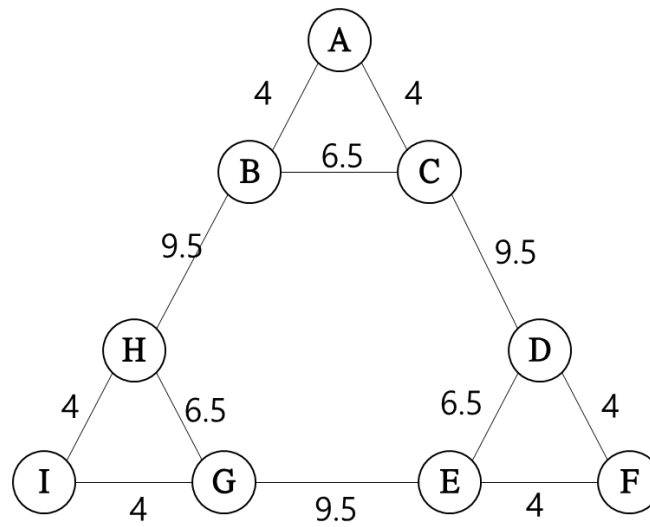
(b) B



Question 2

Using symmetry, the calculations of Question 1 are all you need to compute the betweenness of each edge. Do the calculation. [25 points]

- Solution)



Question 3

For the graph of Question 1, answer the following questions. [30 points]

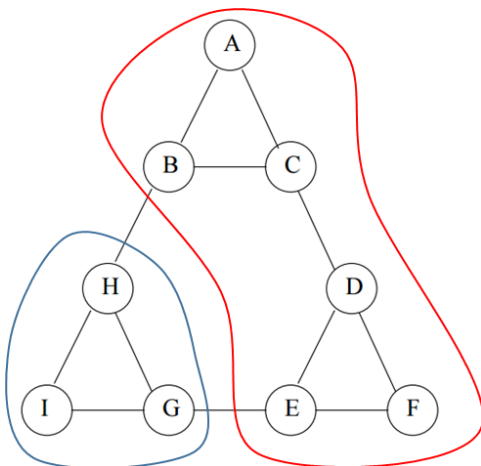
(a) Write down the Laplacian matrix of the graph.

- **Solution :** Laplacian matrix, $L = D - A$ where A is the adjacency matrix and D is the degree matrix.

	A	B	C	D	E	F	G	H	I
A	2	-1	-1	0	0	0	0	0	0
B	-1	3	-1	0	0	0	0	-1	0
C	-1	-1	3	-1	0	0	0	0	0
D	0	0	-1	3	-1	-1	0	0	0
E	0	0	0	-1	3	-1	-1	0	0
F	0	0	0	-1	-1	2	0	0	0
G	0	0	0	0	-1	0	3	-1	-1
H	0	-1	0	0	0	0	-1	3	-1
I	0	0	0	0	0	0	-1	-1	2

(b) For the Laplacian matrix constructed in the above, after computing the second-smallest eigenvalue and its eigenvector, draw two partitions of the graph based on the eigenvector. The splitting point is zero. (Hint: use 'eig' function of MATLAB or OCTAVE)

- There would be multiple second-smallest eigenvalues; thus multiple answers are possible. In that case, just select one of them (i.e., select the first appeared one). If your partition result is based on the second-smallest eigenvalue, then we will give you a full point.
- **Solution:** If an answer is based on the second-smallest eigenvalue and its vector, it is correct.



- **Second eigenvalue: 0.6972**

NODE	EIGENVECTOR	PARTITION
		LABEL(SPLIT AT 0)
A	-0.2677	RED
B	-0.0691	RED
C	-0.2797	RED
D	-0.3073	RED
E	-0.1087	RED
F	-0.3193	RED
G	0.3764	BLUE
H	0.3884	BLUE
I	0.5870	BLUE

Question 4

For the example graph from Figure 2, how many instances of $K_{s,t}$, complete bipartite subgraph, are there for: [20 points]

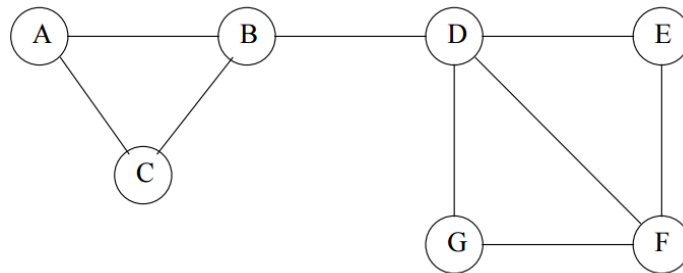


Figure 2. Example graph

(a) $s = 1$ and $t = 3$

• 6

(b) $s = 2$ and $t = 3$

• 0