

Department of Computer Science and Engineering, S V N I T, Surat  
MID-SEMESTER EXAMINATIONS, March 2023  
B. Tech. – III (CSE) – 6<sup>th</sup> Semester  
Course: (CS342) Core Elective (Social Network Analysis)

Date: 6<sup>th</sup> March, 2023

Time: 11:00 to 12:30

Max Marks: 30

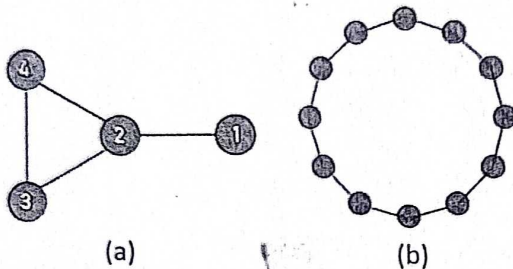
Q.1 Answer the following (Any Six):

[15]

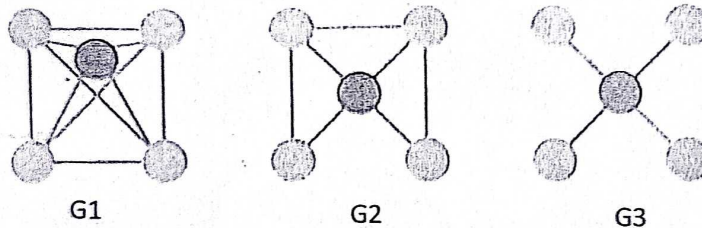
(1) Support or refute the statement with proper justification: "Real networks are really random and unpredictable."

(2) For the adjacency matrix representation of a network  $G(V, E)$ , what are the time complexities; a) to add an edge, b) to remove an edge, c) to add a node and d) to remove a node?

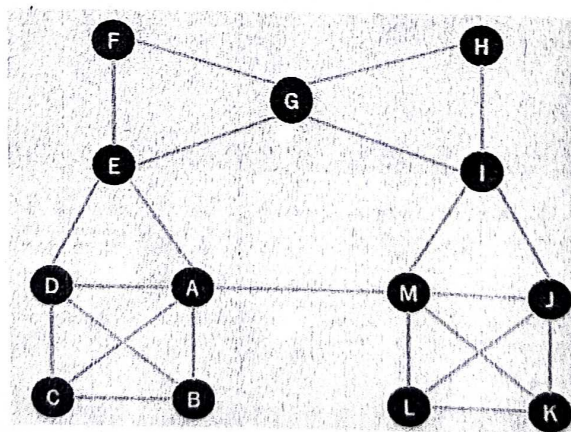
(3) Draw degree distribution graph for the following graphs (a) and (b).



(4) Calculate the clustering co-efficient for the center node in following graphs G1, G2 and G3. What are your observations from the calculated value?



(5) For the following graph, what is the probability that nodes B and M will form a link in future?



(6) List the properties of real-world network.

Which of these properties are satisfied in the below models?

1) Erdos-Renyi random network model

2) Watts-Strogatz model

(7) Barabasi-Albert preferential attachment model

Discuss the Network Schema and Metapath. How these are useful in PathSim algorithm?

Q.2 Answer the following:

(1) In figure 1, a citation network is shown. The edge from node v1 to v2 represents that paper v1 cites paper v2. Two papers are similar if they are cited by similar papers. Calculate the similarity between nodes C and F using the above criteria using SimRank algorithm. Take the value of constant  $C=0.2$ .

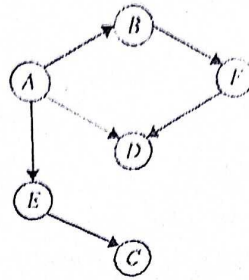


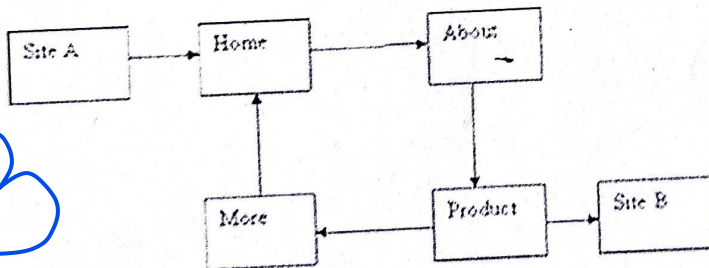
Figure 1

(2) Suggest suitable type of graph that can be used for the following applications. You can consider following types of graph or combination (directed, undirected, weighted, unweighted, signed, unsigned, homogeneous, heterogeneous, unipartite, bipartite).

- 1) Telephone call network (showing frequency of calling from one user to another)
- 2) Email message network (showing email communication between users)
- 3) Film actor collaboration network (showing relationship between actor to movie)
- 4) Academic co authorship network (showing co-author relationship between academicians)
- 5) Connection between houses and utilities (e.g. gas, water, electricity). (showing which utilities are used by houses in city)

(3) Find out the rank of each webpage for the below graph of webpages using PageRank algorithm following the random surfer model. Show computation for only one iteration. Assume following data:

- 1) Initial rank values of all pages as  $1/6$ .
- 2) Damping factor  $d=0.8$





Department of Computer Science and Engineering, S V N I T, Surat  
END-SEMESTER EXAMINATIONS, May 2023  
B. Tech. – III (CSE) – 6<sup>th</sup> Semester  
Course: (CS342) Core Elective (Social Network Analysis)

Date: 1<sup>st</sup> May, 2023

Time: 9:30 to 12:30

Max Marks: 50

Q.1 Answer the following:

[15]

(1) In order to obtain diversity in the result, one can consider running clustering algorithm at all nodes and returning centroids of clusters. Why do we need to use DivRank?

(2) Explain pros and cons of following community detection algorithms used to detect overlapping communities.

(a) Clique Percolation

(b) Link Partition

(3) Differentiate Graph Partitioning and Community Detection.

(4) How do you identify an edge acting as a local bridge in the network? Explain with example.

(5) Explain the idea behind BigClam algorithm. How it is different from Hierarchical clustering and Greedy algorithms for community detection?

Q.2 Answer the following:

[25]

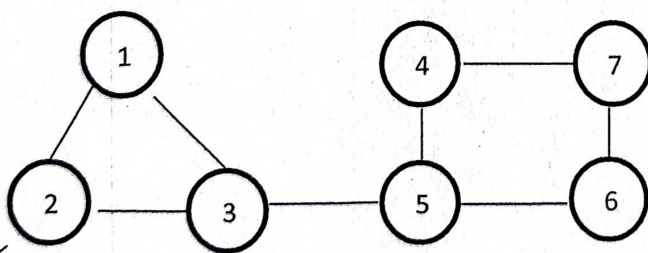
(1) Consider the following graph and Construct the similarity matrix for Ravasz algorithm for Agglomerative Clustering.

[5]

Use the following similarity measure as discussed in the class:

$$x_{ij}^0 = \frac{j(i, j)}{\min(k_i, k_j) + 1 - \theta(A_{ij})}$$

- $\theta(x)$  is the Heaviside step function, which is zero for  $x \leq 0$  and one for  $x > 0$ ;
- $J(i, j)$  is the number of common neighbors of node  $i$  and  $j$ , to which we add one (+1) if there is a direct link between  $i$  and  $j$ ;
- $\min(k_i, k_j)$  is the smaller of the degrees  $k_i$  and  $k_j$



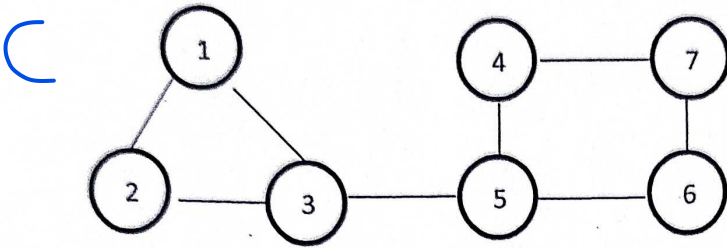
(2) A bibliographic network is a typical heterogeneous network, containing objects from four types of entities: papers (P), venues (i.e., conferences/journals) (C), authors (A), and terms (T). For each paper  $p \in P$ , it has links to a set of authors, a venue, a set of words as terms in the title, a set of citing papers, and a set of cited papers, and the link types are defined by these relations.

[5]

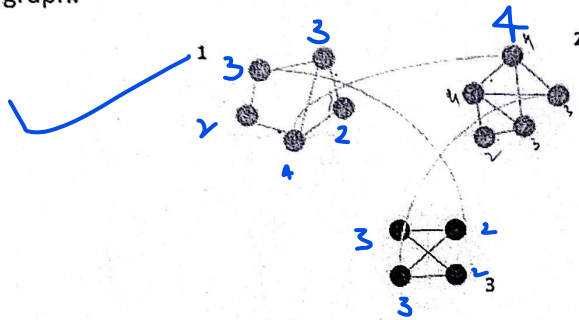
Define the network schema and identify Meta Paths for this network.

$$M = \left( \frac{L_c}{2L} - \left( \frac{K_c}{2L} \right)^2 \right)$$

- (3) Detect the communities in the following social network using Louvain algorithm. Show all steps. Take node 1 as the seed node. [10]



- (4) Consider the graph shown below. The graph is partitioned into three disjoint communities. Calculate modularity for the graph. [5]



Q.2 Answer the following:

- (1) Explain Graph Representation Learning pipeline in detail with its components [10]  
 (2) Explain in detail: Multiple choice decision-based cascade model with example. Which strategy would be adopted by node u and v? Consider the payoff  $a=5, b=3, c=1$ . [2.5]



- (3) Explain SIR model. What is the rate of change of SIR? [2.5]

$$\begin{aligned} \frac{ds}{dt} &= -\beta SI \\ \frac{di}{dt} &= \beta SI - \delta I \\ \frac{dr}{dt} &= \delta I \end{aligned}$$

Cheney

$$M = \left( \frac{L_c}{2L} - \left( \frac{K_c}{2L} \right)^2 \right)$$

$$\begin{aligned} n_c &= 3 \\ L &= \\ L_c &= \\ K_c &= \end{aligned}$$