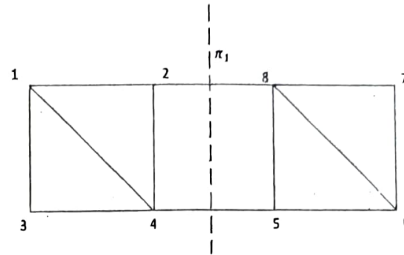


Note: Marks are indicated against each question. Parts of a question must be answered together.

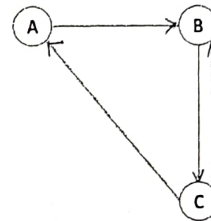
Q1. Attempt any five questions out of the following: [5 * 6marks =30]

a) What is minimum cut problem for finding communities? Compute the ratio-cut and normalized ratio-cut for the partition π_1 in the below graph.



b) Write short note on the paradigm of Web Analytics 2.0.

c) Compute the page rank of the webpages A, B, C and rank them in decreasing order. ($d=0.85, \epsilon=0.09$)



- d) Discuss any three types of centrality with formulae/examples.
- e) Explain clustering coefficient of a node and of entire graph with an example.
- f) Differentiate between paid search and organic search.

Q2. Attempt any two questions out of the following: [2* 10marks=20]

- a) Explain the bottom-up approaches for identifying subgroups.
- b) Discuss the structural and regular equivalence by taking an example. Give measure to compute the degree of equivalence for structural equivalence for finding equivalence sets.
- c) Consider the following rating matrix for a collaborative filtering problem where Dave is the *active* user. Compute neighbourhood ($k=2$) for Dave using pearson correlation as the similarity metric. What is Resnick's prediction formula and briefly describe how can it be used to make item recommendations to Dave.

	I1	I2	I3	I4	I5
Alex	-	-	3	2	5
Bruce	4	1	1	5	4
Carolina	-	-	5	3	4
Dave	2	?	4	?	?
Evelyn	2	4	4	3	-

Q3. Attempt any five questions out of the following: [5 * 2marks =10]

- a) An advertising campaign costs \$2000 and generates 40,000 impressions. Compute CPM.
- b) Define the following terms: i) Closed Walk ii) Eccentricity
- c) Briefly describe structural holes with an example.
- d) Explain in brief the clique percolation method to identify clusters.
- e) Discuss the implications of engagement rate on bounce rate with a suitable example.
- f) What are two mode networks?
