MITE
Invent Solutions
MANGALORE INSTITUTE OF TECHNOLOGY & ENGINEERING
(A Unit of Rajalaxmi Education Trust®, Mangalore)
Autonomous Institute affiliated to VTU, Belagavi, Approved by AICTE, New Delhi
Accredited by NAAC with A+ Grade & ISO 9001:2015 Certified Institution

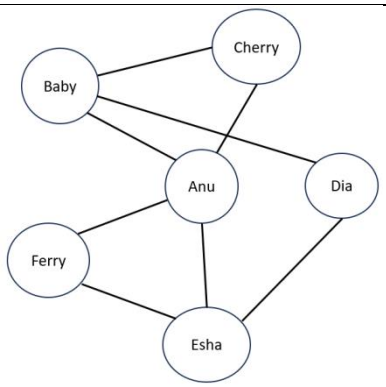
Model Question Paper

Sixth Semester BE Degree Examination

Social Network Analysis

Time: 3 Hours(180 Minutes) **Max. Marks: 100**

Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.
2. M: Marks, L: RBT (Revised Bloom's Taxonomy) level, C: Course outcomes.

		Module -1	M	L	C
Q1	a.	Explain Social Network Analysis (SNA) with an example.	06	L2	CO1
	b.	 *Figure 1: Friends Graph* With reference to the example friends graph given in figure 1, depicting the social network connections among friends, construct a adjacency matrix and adjacency list to determine the most influenced person and justify your answer.	07	L2	CO1
	c.	Discuss in detail, different levels of Social Network Analysis.	07	L2	CO1
		OR			
Q2	a.	With an example, explain different types of Clustering coefficients and discuss how it is used in Social Network Analysis.	06	L2	CO1
	b.	By considering an e-commerce example "**user-product network**", explain how a Bipartite network can always induce a Unipartite network.	07	L2	CO1
	c.	Draw an example graph, depicting the Facebook friends connection, explain the application of Social Network Analysis in this context to determine friends and follow recommendation.	07	L2	CO1
		Module- 2			
Q3	a.	With reference to example dataset given in Table 1 *Table 1: Example_Dataset* User-i | User-j | No of comments on a post Alice | Bob | 5 Alice | Charlie | 3 Bob | David | 4 Charlie | David | 2 David | Eva | 6 Write a R program to convert the given tabular data into a graph using the igraph package. Generate appropriate network visualization. Enhance the visualization by applying effective design principles: i. Assign different colors to nodes based on interaction intensity (e.g., high vs low degree). ii. Adjust node size proportional to degree centrality.	08	L3	CO2

		iii. Customize font style, size, and color of labels for better readability. Illustrate how above network visualization elements (color, size, font) improve the analysis and understanding of above social network situation.			
	b.	Compare and illustrate, with a suitable example (e.g., a friendship network), the advantages of the following graph layout techniques used in social network analysis: • Random layout and • Circle layout Justify your answer which layout is most suitable for community detection and large-scale network.	06	L3	CO2
	c.	*Table 2 : Example_Dataset* With an R Program, illustrate the transformation of the employee data given in table 2 into an edge list & to construct and visualize the management hierarchy using a dendrogram & justify why a tree structure is appropriate.	06	L3	CO2

EmployeeId	FirstName	Department	ReportsTo
1	Arun	Management	NA
2	Bhavna	Sales	1
3	Kiran	HR	1
4	Meena	Sales	2
5	Rakesh	Sales	2
6	Suresh	HR	3
7	Divya	IT	3
8	Naveen	IT	7
9	Pooja	IT	7
10	Anil	Finance	1

OR

Q4	a.	*Table 3: Example_Dataset* With reference to example dataset giving in table 3, Write an R program to convert the given tabular data into a graph using the igraph package. Generate an appropriate network visualization. Enhance the visualization by applying effective design principles: i. Assign different border colors to vertex (e.g., based on incoming and outgoing degrees) ii. Adjust edge width, color, arrow size and arrow mode based on the weights between vertices. iii. Customize line type of edge and amount of curvature to the edge for better visibility among connected edges. Illustrate how above network visualization elements (edge color, edge type and arrow type) improve the analysis and understanding of above social network scenario.	08	L3	CO2
	b.	Compare and illustrate, with a suitable example (e.g., communication network), the advantages of the following graph layout techniques used in social network analysis: • Force-directed layout and • Grid / hierarchical layout justify your answer which layout is most suitable for large scale network, structured or hierarchical networks.	06	L3	CO2
	c.	With reference relational tables given in Figure 2: Customers(CustomerId, FirstName, LastName), Invoices(InvoiceId, CustomerId), Items(InvoiceId, TrackId)	06	L3	CO2

From	To	Distance
Agra	Bangalore	6
Agra	Chennai	4
Bangalore	Delhi	5
Chennai	Delhi	3
Delhi	Ernakulam	7

CustomerId	FirstName	LastName
1	Rahul	Sharma
2	Anjali	Verma
3	Kiran	Rao
4	Sneha	Iyer
5	Arjun	Mehta

InvoiceId	CustomerId
101	1
102	2
103	3
104	4
105	5

InvoiceId	TrackId
101	201
101	202
102	202
103	203
104	201

Figure 2: Example_Dataset

		a) Write R code to construct a dataset of distinct *(Customer, TrackId)* pairs suitable for network analysis. and to generate and visualize the corresponding **bipartite network**, clearly distinguishing node types. Illustrate with an example, the impact of applying SNA in this context with respect to identifying similar customers, recommendation opportunities, and community structures.			
		Module – 3			
Q5	a.	*Table 4 : Example_Dataset* From City / To City / Weight(Distance) 9 10 4 10 11 1 11 12 1 10 12 1 9 13 3 13 14 2 9 8 2 9 7 3 8 7 1 4 6 1 4 7 2 4 8 3 6 7 2 4 1 1 4 2 1 4 3 1 4 5 2 1 2 1 Draw a graph using dataset given in table 4 and also Write an R program to convert the given tabular data into a graph using the igraph package and to do the following tasks i. Create weighted graph and using R function find the **all simple paths** from city-9 to city-4 ii. On obtained weighted graph and using R function find the **all shortest paths** from city-9 to city-4 Illustrate how tree traversal improves the understanding of social network scenario in the context of a city network.	08	L3	CO3
	b.	Consider two social network graphs G1 and G2 with the following properties: • G1: Diameter = 5, Average distance = 2.38, Density = 0.2 • G2: Diameter = 4, Average distance = 2.49, Density = 0.2 Answer the following: 1. Apply the concept of diameter to determine which network appears more close-knit and justify your answer and Apply the concept of average distance and compare the closeness of the two networks. 2. Interpret the role of density in both networks and explain why they are considered structurally similar or different. Justify your answer on SNA reasoning on why multiple metrics (diameter, average distance, density) are required instead of relying on a single measure.	06	L3	CO3
	c.	*Table 5 : Example_Edge_list* from / to 1 2 1 3 2 4 3 4 4 5 With reference to the graph edge list given in table 5 Write an R program to: 1. Construct the graph from the given dataset. 2. Compute the **distance matrix** using appropriate functions. 3. Identify the **closest and farthest node pairs**. Apply SNA concepts to explain how distance influences Social closeness & Likelihood of communication	06	L3	CO3
		OR			
Q6	a.	With reference to the tweeter example friends graph given in Figure 3, representing the social network connections among friends, and analyze the tweeter network for following graph components: a walk, a path, a cycle and illustrate your answer how above components can be applied for social network analysis in this context.	08	L3	CO3

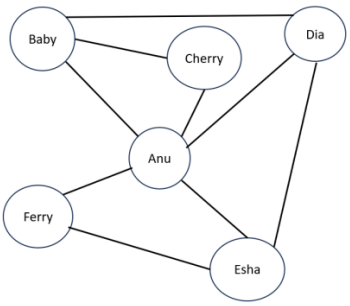


Figure 3: Example_Graph

b. *Table 6:Example_graph*

from	to
A	B
B	C
C	D
D	E
B	E

With reference to the edgelist for graph given table 6, Write an R program to:

1. Construct the graph from the dataset & Compute the **distance matrix** using appropriate functions.
2. Identify a pair of nodes with **maximum distance**.
3. Select a group of 3 nodes that are relatively **far from each other**.

Illustrate how above metrics can applied for Social Network Analysis in this context

06 | L3 | CO3

c. *Table 7 : Example_Dataset*

from	to
Monica	Chandler
Monica	Rachel
Rachel	Ross
Chandler	Joey
Joey	Phoebe
Ross	Phoebe
Janice	Chandler
Mr Bing	Chandler
Mrs Bing	Ross

With reference to the following table 7, Write an R program to:

1. Find **all simple paths** from *Janice* to *Mrs Bing*.
2. Compute the **shortest path distance** between them.

Interpret the result using SNA concepts in terms of: Familiarity between characters, Likelihood of information flow

06 | L3 | CO3

Module – 4

Q7

a. Apply appropriate methods to compute the degree centrality of vertices in an undirected graph. Using the graph shown in Figure 4, calculate the degree centrality of vertices 9, 10, and 11. Further, demonstrate how degree centrality can be used to analyze the importance of individuals in an organizational Facebook network.

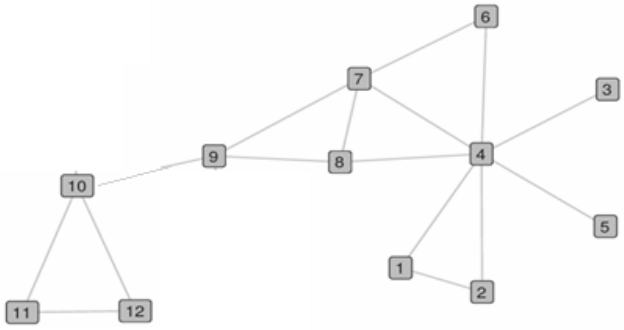


Figure 4 : Example Graph

08 | L3 | CO4

b. Consider the following graphs:
Graph 1 (Undirected): Vertices {A, B, C, D}, Edges {A–B, B–C, C–D}
Graph 2 (Directed): Vertices {P, Q, R}, Edges {P → Q, Q → R}
Graph 3 (Directed): Vertices {X, Y, Z}, Edges {X → Y, Y → Z, Z → X}

06 | L3 | CO4

		i. Apply the concept of connectivity to determine whether Graph 1 is connected. ii. For Graph 2, determine whether it is weakly connected or strongly connected by analysing reachability between vertices. iii. For Graph 3, determine whether it is strongly connected and justify your answer. Based on the above examples, demonstrate the differences between connected, weakly connected, and strongly connected graphs.			
	c	With reference to the facebook example dataset given in table 8 & table 9 ***Table 8: Example_Dataset_User_Type*** ***Table 9 : Example_Dataset_Class*** Write an R program to load a 'schoolfriends' graph dataset representing a Facebook friends network and perform the following operations: i. Construct an undirected graph and remove isolated vertices (isolates). Apply suitable R functions to detect communities and cliques within the network. ii. Using the results obtained, determine the largest clique and demonstrate how cliques are distributed across different classes/groups in the network	06	L3	CO4
		OR			
	a.	Consider the following people social network represented as an undirected graph: Vertices: {Abhi, Bhanu, Cia, Dhanush, Elfrid, Fathima} Edges:{(Abhi – Bhanu), (Abhi –Cia), (Bhanu –Cia), (Bhanu – Dhanush), (Cia–E), (Dhanush – Elfrid), (Dhanush – Fathima), (Elfrid – Fathima)} i. Apply the centrality measures—degree, closeness to compute the centrality values for each vertex. ii. Identify the individuals with the highest centrality under each measure. iii. Demonstrate how these results can be used to support real-world decisions such as identifying key influencers, improving communication efficiency, and selecting leaders in the network.	08	L3	CO4
Q8	b.	Consider the following graph representing a network: Vertices: {A, B, C, D, E, F, G, H} Edges: {(A–B), (A–C), (B–C), (B–D), (C–D), (E–F), (E–G), (F–G), (F–H), (G–H), (D–E)} i. Apply the concept of community detection to identify communities (groups of closely connected vertices) in the graph. ii. Justify the identified communities based on the density of connections within and between groups. iii. Demonstrate the objective of community detection by explaining how the identified communities help in understanding the structure and relationships within the social network scenario.	06	L3	CO4
	c.	With reference to example dataset given in table 10 representing votes from **Wikipedia members** to other Wikipedia members to nominate **from member to administrator** status.	06	L3	CO4

From	To	Type
11	88	facebook
11	99	facebook
22	11	facebook
33	55	reported
33	88	reported
44	22	facebook
44	66	facebook
44	88	facebook
55	11	reported
66	55	facebook
77	88	facebook
88	77	reported
88	44	reported
88	11	reported
99	11	reported

id	class	gender
11	AIML	F
22	ECE	F
33	CSE	F
44	ISE	M
55	MECH	M
66	MECH	F
77	CIVIL	M
88	AIML	F
99	ECE	F

Table 10 : Example_Dataset

From	To
3	39
3	54
3	108
6	54
6	39
8	178
8	39
8	214
12	54
12	39

Write the R program to load the wikivote edgelist & Create a directed graph & to determine the following

i. How many **weakly** connected components there are in this graph. How large is the largest component?

ii. how many **strongly** connected components there are in this graph. How large is the largest component?

Illustrate how you would analyze and interpret the difference between your results for Questions i and ii with respect to Social Network Analysis.

Module – 5

Q9

a.

Vertices	
Node	**Class**
1	A
2	A
3	B
4	B
5	A
6	B

Edges	
from	**to**
1	2
1	3
2	5
3	4
4	6
5	6

Figure 5: Example_Dataset

With reference to the following tables given in Figure 5 providing information about graph vertices & edge.

Construct the graph from the given dataset and classify all edges into same-class and different-class connections. Using this classification, illustrate the concept of the assortativity coefficient in Social Network Analysis, and determine whether the network is assortative, disassortative, or neutral based on the edge distribution.

10 | L3 | CO4

b

Design a labelled-property graph schema for a movie database by identifying:
Node types (e.g., Person, Movie), Relationship types (e.g., ACTED_IN, DIRECTED) and identify suitable Key properties for nodes and relationships.
And Construct suitable Cypher queries to:

a) Retrieve all actors who acted in movies directed by a "Edward"
b) Display actor name, movie title, and release year

Justify your answers on how such a graph database structure supports efficient relationship queries compared to traditional databases, particularly in the context of Social Network Analysis.

10 | L3 | CO5

OR

Q10

a. For the graph given in question number 7a (Figure 4), Compute Jacquard, dice & inverse log-weighted similarity coefficient and illustrate its significance on determining vertex similarity in social network analysis.

10 | L3 | CO4

b.

Conceptual Sample Dataset for Online Learning Platform			
Users:	U1	U2	U3
Courses:	C1 (Data Science)	C2 (Machine Learning)	
Interactions:	U1 Enrolled in C1		
	U2 enrolled in C1 and C2		
	U3 enrolled in C2		
	U1 interacts with U2		
	U2 interacts with U3		

Figure 6 : Conceptual Sample Dataset

With reference to the Online Learning Platform conceptual sample dataset given in Figure 6, Design a labelled-property graph schema by identifying: Node types (e.g., User, Course), Relationship types (e.g., ENROLLED_IN, INTERACTS_WITH), Properties for nodes (e.g., userId, courseName) and edges (e.g., interactionType)

Represent the dataset as a graph structure using nodes and relationships. & write Cypher queries to:

a) Retrieve all users enrolled in a specific course
b) Find users who interact with others enrolled in the same course

Justify your answer on how this graph model supports efficient relationship analysis in Social Network Analysis.

10 | L3 | CO5

-*****-