



المادة: Social Network Analysis

المدة: ساعتين

الأستاذ: د. حسين هزيمة

المرحلة: الإجازة

السنة المنهجية: الثالثة

الاختصاص: علم البيانات - Data Science

Documents are NOT authorized

Part 1: (20 pts) multiple-choice questions:

1. What does a node represent in Social Network Analysis?

- A. A relationship
- B. A graph
- C. An actor or entity
- D. A matrix

2. In a directed graph, the adjacency matrix is usually:

- A. Symmetric
- B. Diagonal
- C. Non-symmetric
- D. Empty

3. Which centrality measure counts the number of direct connections of a node?

- A. Closeness centrality
- B. Betweenness centrality
- C. Degree centrality
- D. Eigenvector centrality

4. In an undirected graph, the adjacency matrix becomes:

- A. Sparse
- B. Symmetric
- C. Directed
- D. Weighted

5. What is an ego network?

- A. A network without nodes
- B. A complete graph
- C. A network centered around one actor
- D. A disconnected graph

6. Which of the following is NOT a possible edge weight?

- A. Frequency of interaction

- B. Distance
- C. Relationship strength
- D. Node color

7. Betweenness centrality measures:

- A. Number of neighbors
- B. Fraction of shortest paths through a node
- C. Density of graph
- D. Node labels

8. Closeness centrality identifies nodes that:

- A. Have the most colors
- B. Spread information efficiently
- C. Have no neighbors
- D. Form cliques

9. A component in a graph is:

- A. A weighted edge
- B. A disconnected node
- C. A group of connected nodes
- D. A matrix representation

10. Eigenvector centrality gives high scores to nodes connected to:

- A. Isolated nodes
- B. Random nodes
- C. Important nodes
- D. Weak edges

11. Which graph property describes six degrees of separation?

- A. Dense graph
- B. Diameter
- C. Clique
- D. Directed graph

12. A clique is:

- A. A disconnected subgraph
- B. A complete subgraph
- C. A weighted graph
- D. A sparse graph

13. Which community detection method uses cliques as a core?

- A. BFS
- B. DFS
- C. Clique Percolation Method
- D. Dijkstra Algorithm

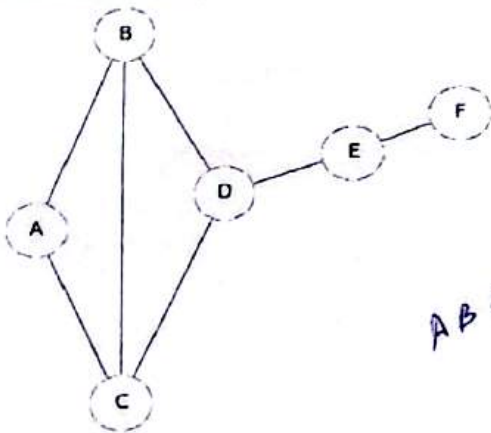
14. In CPM, two cliques are adjacent if they share:
- A. One node
 - B. No nodes
 - C. $k-1$ nodes
 - D. All nodes
15. Heavy-tailed degree distribution means:
- A. All nodes have equal degree
 - B. Most nodes have high degree
 - C. Few hubs and many low-degree nodes
 - D. No hubs exist
16. The Katz index is used for:
- A. Graph coloring
 - B. Link prediction
 - C. Community deletion
 - D. Matrix inversion
17. Jaccard coefficient measures:
- A. Graph diameter
 - B. Neighbor overlap
 - C. Number of edges
 - D. Centrality rank
18. Weisfeiler-Lehman kernel uses:
- A. Color refinement
 - B. Edge deletion
 - C. BFS search
 - D. Node removal
19. Graphlet kernels represent graphs as:
- A. Bag-of-graphlets
 - B. Trees
 - C. Random walks
 - D. Shortest paths only
20. Which feature is useful for predicting influential users?
- A. Random colors
 - B. Importance-based features
 - C. Sparse matrices
 - D. Edge weights only

Part 2: (20 pts) True/False questions:

1. In undirected graphs, adjacency matrices are symmetric. T
2. Degree centrality only applies to directed graphs. _____
3. Closeness centrality depends on shortest path distances. _____
4. Betweenness centrality requires a fully connected graph. _____
5. Eigenvector centrality considers the importance of neighboring nodes. _____
6. A clique is a sparse subgraph. _____
7. Ego networks focus on one central actor and its neighbors. _____
8. Edge weights can represent communication frequency. _____
9. Components are groups of disconnected nodes only. _____
10. Heavy-tailed distributions contain many hubs. _____
11. CPM stands for Clique Percolation Method. _____
12. Dense graphs contain relatively many edges. _____
13. Jaccard coefficient is a local neighborhood overlap metric. _____
14. Katz index only counts paths of length 1. _____
15. Graphlet kernels are computationally cheap for large graphs. _____
16. WL kernel uses iterative color refinement. _____
17. Distance-based features capture neighborhood overlap perfectly. _____
18. Degree centrality counts neighboring nodes. _____
19. Small-world phenomenon refers to large graph diameters. _____
20. Social Network Analysis combines sociology and graph theory. _____

Part 3: (60 pts) Problem solving questions:

Problem 1: (30 pts) Betweenness Centrality

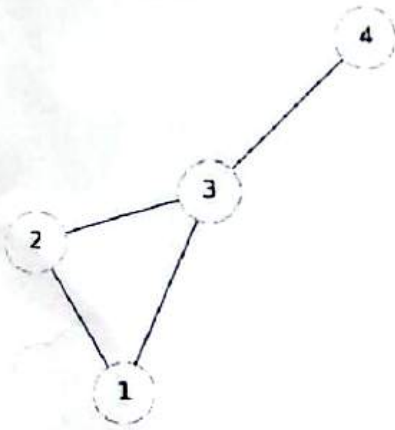


ABCDEF

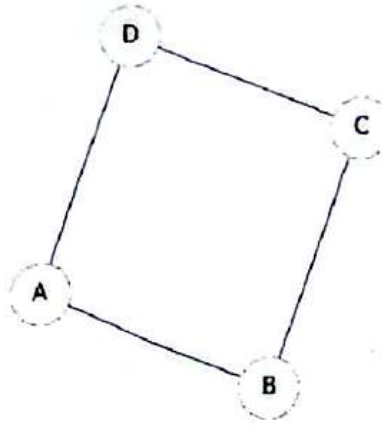
1. List all shortest paths between nodes A and F.
2. Calculate the betweenness centrality contribution for node D from paths between A and F.
3. Determine which node has the highest betweenness centrality and explain why.
4. Explain why nodes with high betweenness centrality are important in social networks.

Problem 2: (30 pts) Graphlet Kernels

Graph G1



Graph G2



1. Count the number of triangle graphlets in G1 and G2.
2. Count the number of path graphlets in G1 and G2.
3. Construct the graphlet count vectors for both graphs in the form [#Triangles, #Paths].
4. Compute the graphlet kernel value using the inner product between the two vectors.
5. Explain what the kernel value indicates about the similarity between the two graphs.

Good Work

Appendix

$cc(v) = 2N_v / K_v(K_v - 1)$	$s = C(n, K)$ where $k = n(n-1)/2$
$C_{D_i} = \frac{\sum_{j=1}^n a_{ij}}{n-1}$	$C(x) = \frac{N-1}{\sum_y d(y, x)}$
$C_{btw}(v) = \sum_{s,t \in N} \frac{\sigma_{s,t}(v)}{\sigma_{s,t}}$	$c^{(k+1)}(v) = \text{HASH}(\{c^{(k)}(v), \{c^{(k)}(u)\}_{u \in N-v}\})$
$a \cdot b = \sum_{i=1}^n a_i b_i = a_1 b_1 + a_2 b_2 + \dots + a_n b_n$	$K(G, G') = \Phi(G)^T \Phi(G')$
$K(G, G') = f_G^T f_{G'}$	