Social Network Analysis
(CSEN 6159)

Time Allotted : 3 hrs   Full Marks : 70

Figures out of the right margin indicate full marks.

Candidates are required to answer Group A and
any 5 (five) from Group B to E, taking at least one from each group.

Candidates are required to give answer in their own words as far as practicable.

Group – A
(Multiple Choice Type Questions)

1. Choose the correct alternatives for the following: 10 x 1=10

(i) Results from the experiments designed by Milgram support the theory of
  (a) Diffusion of information   (b) Six degrees of separation
  (c) Preferential attachment   (d) None of these.

(ii) Degree distribution of a social network usually follows
  (a) Gaussian distribution   (b) Power-law distribution
  (c) Exponential distribution   (d) Poisson distribution.

(iii) Which is not TRUE about an Erdos-Renyi network?
  (a) It may follow a $G(n, m)$ model   (b) It may follow a $G(n, p)$ model
  (c) Degree distribution is binomial   (d) It is a scale-free network.

(iv) Features of an SIS model can be observed in
  (a) common cold   (b)   swine flu
  (c) bird flu   (d)   malaria.

(v) Closeness value should lie within the range,
  (a) 0 to 1 (both inclusive)   (b) -1 to 0 (both inclusive)
  (c) -0.5 to 0.5 (both inclusive)   (d) -1 to 1 (both inclusive) .

(vi) Which of the following is an overlapping community detection algorithm?
  (a) Louvain   (b) Girvan-Newman
  (c) Clique percolation   (d) Clauset-Newman-Moore (CNM).

(vii) Propagation probability is a parameter in
  (a) SIS model   (b) SIRS model
  (c) Independent Cascade model   (d) Linear Threshold model.

(viii) Diameter of a network is the
  (a) shortest of all shortest paths (b) longest of all longest paths
  (c) longest of all shortest paths   (d) shortest of all longest paths.

(ix) Betweenness can be calculated in
(a) $O(n)$ (b) $O(n^2)$ (c) $O(n\log n)$ (d) $O(n^3)$.

(x) Hub and authority values can be calculated by,
(a) HITS   (b) PageRank
(c) Bonacich's power centrality   (d) Betweenness.

Group – B

2.(a) Define Bonacich's power centrality and explain how its mathematical representation can accommodate two contradictory ideas of how powerful a node is. Describe the mathematical equation to find power using Bonacich's power centrality.

(b) What is degree centrality? "Degree distribution in a social network usually follows power-law distribution" – what is the significance of this statement?

$$(2 + 2+ 3) + (2 + 3) = 12$$

3.(a) Explain the concept of "Preferential attachment". Prove that preferential attachment leads to generation of a scale-free network.

(b) Define modularity and explain it with necessary mathematical equations.

$$(2 + 6) + 4 = 12$$

Group – C

4.(a) What is PageRank? How is i) Dead-end and ii) Spider trap, related to PageRank? How can they be avoided?

(b) Explain the concepts of i) betweenness, ii) local clustering co-efficient, with one example of each.

$$(2 + 2 \times 2.5) + (3 + 2) = 12$$

5.(a) Find out speed of an epidemic using Fisher's equation in an SI model. Start by writing the differential equations for SI model using t and x.

(b) What is conductance? How is it related to detection of communities in social networks?
$$7 + (2 + 3) = 12$$

Group – D

6.(a) What is the reproductive number in epidemiology? How can we categorize the nature of an epidemic in terms of reproductive number?

(b) Describe Louvain method for detecting communities in social networks. State all the stages and describe them separately. Clearly mention the equation which helps calculating change in modularity. Comment on the time complexity of the method.
$$4 + (5 + 2 + 1) = 12$$

7.(a) Write the pseudo-code/ C program to implement the independent cascade model. Assume that the graph information is already available in an adjacency list

(b) What is power iteration method in PageRank? Take a sample directed graph of 5 nodes and at least 8 edges and calculate its PageRank by the power iteration method (stop after 4 iterations).

$$6 + 6 = 12$$

Group – E

8.(a) Describe the SIR model and find out the relationship between $S(t)$ and $I(t)$.

(b) Describe Barabasi-Albert network model and its important features.
$$7 + 5 = 12$$

9.(a) Explain Girvan-Newman (GN) algorithm for detection of community structures in social networks. Comment on the time complexity of the algorithm. What are the limitations of GN algorithm?

(b) Describe small-world network model as described by Watts-Strogatz.
$$7 + 5 = 12$$