

**SOCIAL NETWORK ANALYSIS
(CSEN 4245)**

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 alternative for the following: **10 × 1 = 10**
- (i) "Six degrees of separation" theory is associated with
(a) complex network (b) random network
(c) small-world network (d) none of these.
- (ii) Degree distribution of a random Erdos-Renyi network usually follows -
(a) Pareto distribution (b) Zipf distribution
(c) Weibull distribution (d) Poisson distribution.
- (iii) Power-law degree distribution can be observed in which type of network?
(a) Complex network (b) Random network
(c) Small-world network (d) Scale-free network.
- (iv) The normalization factor for calculation of betweenness of a node is given by,
(a) $n(n-1)$ (b) $n(n-1)/2$
(c) $(n-1)(n-2)/2$ (d) n^2 .
- (v) Modularity value can range within
(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) Maximal clique finding algorithm is used by which of the following community detection technique?
(a) CONGA (b) Girvan-Newman
(c) Clique percolation (d) None of these.

- (vii) Value of clustering co-efficient may range from
 (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).
- (viii) Which of the following is not a compartmental model for epidemics?
 (a) SIR (b) SIS
 (c) SISR (d) SIRS.
- (ix) In small-world network, approximate maximum size of the diameter is
 (a) 6 (b) 12
 (c) 9 (d) dependent on the graph.
- (x) Fisher's equation is used to calculate,
 (a) speed of an epidemic (b) no. of infected
 (c) no. of recovered (d) none of these.

Group - B

2. (a) What is PageRank? What 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 each.
- (1 + 2 + 2 + 1) + (2 × 3) = 12**
3. (a) In a scale-free social network, what is the probability that the degree of a randomly picked node will be k? Explain how it depends on k. Draw the k vs p(k) plot and comment about its nature.
- (b) Explain the concept of global clustering co-efficient. What will be the global clustering coefficient of the minimum spanning tree of any graph G?
- (2 + 3 + 3) + (2 + 2) = 12**

Group - C

4. (a) Describe the features of a Small-world network model. Deduce the expressions for the three salient features describing this model.
- (b) Calculate the average shortest path length for an ER network with 1 billion nodes and an average degree of one thousand.
- 8 + 4 = 12**
5. (a) Explain the concept of "Preferential attachment". Prove that preferential attachment leads to generation of a scale-free network.

- (b) Explain why adding random shortcuts to a uniform network will decrease its clustering co-efficient.

$$(2 + 7) + 3 = 12$$

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) Explain the nature of the infected compartment at time= t ($I(t)$) in terms of the total population (N) and the infected compartment at time= 0 ($I(0)$) along with the other model parameters in SI model (S should not be included). Comment on the nature of the curve.

$$4 + (7 + 1) = 12$$

7. (a) Describe independent cascade model and linear threshold model for the diffusion of information in social networks.

- (b) Briefly describe the SIR model. Write down all the differential equations that describe this model and derive the relationships between the compartments.

$$5 + (1 + 6) = 12$$

Group - E

8. (a) 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.

- (b) What is label propagation algorithm (LPA)? What is the main advantage of using LPA?

$$(5 + 2 + 1) + 4 = 12$$

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) What is modularity in networks? Explain the mathematical definition of modularity and how suitable it is to be used as a goodness measure for the detection of community structures in a social network.

$$7 + 5 = 12$$

