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
<u>any 5 (five)</u> from Group B to E, taking <u>at least one</u> 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)   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) 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) Features of an SIS model can be observed in
         (a) common cold   (b) swine flu
         (c) bird flu   (d) malaria.

   (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 techniques ?
         (a) CONGA   (b) Girvan-Newman
         (c) Clique percolation   (d) None of these.

   (vii) Propagation probability is a model parameter in which of the following models?
         (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)   In small-world network, approximate maximum size of the diameter is
         (a) 6   (b) 12   (c) 9   (d) dependent on the graph.

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) Calculate the degree distribution of the following graph and plot it (graph paper not needed).
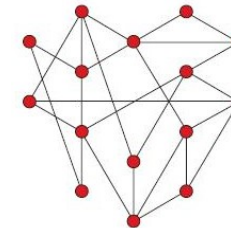
   (b) Define hubs and authorities. Describe HITS algorithm for calculating hub and authority value of nodes.

                                        7 + (2 + 3) = 12

Group – C

4.(a) Show that the degree distribution of ER network will follow Poisson distribution, when N is very large.

   (b) Find out the clustering co-efficient for a uniform triangular lattice and a uniform network where nodes are connected as a ring and each node is connected to its c- nearest neighbours, where c is an even number.

                                        6 + (2 + 4) = 12

5.   Derive and describe the equations for expressing degree distribution, clustering co-efficient and average shortest path length for a small-world network as described by Watts-Strogatz. Explain for every feature if it succeeds (or fails) to replicate the behaviour of a social network.

(3+ 5 + 2 + 2) = 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 $I(t)$ at time=t in terms of the total population (N) and the infected compartment $I(0)$ at time=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)   Write the pseudo-code/ C program to implement the independent cascade model. Assume that the graph information is already available in an adjacency list, with the propagation probability.

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

6 + 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