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) A community is defined as a set of nodes that are
(a) More densely connected to each other
(b) Connected to each other
(c) Connected to all in the network
(d) Densely connected to other community

(ii) A measure of how long it will take information to spread from a given vertex to other reachable vertices in the network is named as
(a) Degree Centrality (b) Betweenness Centrality
(c) Closeness Centrality (d) Eigen Vector Centrality.

(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) Average degree of a random network is denoted by
(a) P(N-1) (b) N(N-1)/2 (c) $N^2/2$ (d) 2^{n-2}

(v) Watts and Strogatz model possesses
(a) Short average path length (b) High clustering
(c) Either (a) or (b) (d) Both (a) and (b).

(vi) The phenomenon "In real networks new nodes tend to link to the more connected nodes" is referred as
(a) Growth of the network (b) High clustering
(c) Preferential Attachment (d) Obeying the Power law.

(vii) Disease state "reinfection" is first introduced in
(a) SIS (b) SIR (c) SIRS (d) Both SIR and SIRS.

(viii) Which of the following is not a compartmental model for epidemics?
(a) SIR (b) SIS (c) SISR (d) SIRS.

(ix) Networks with power-law degree distribution are often called
(a) Real Networks (b) Scale free network
(c) E-R Network (d) Small-scale Network.

(x) "Six degrees of separation" theory is associated with
(a) Complex network (b) Random network
(c) Small-world network (d) None of these.

Group – B

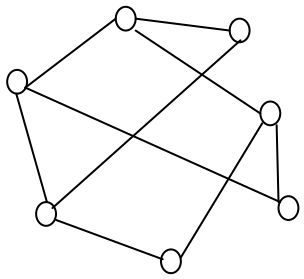
2. (a) What is PageRank? Explain two potential problems that we may face during the calculation of PageRank.

(b) If PR(A) represents the PageRank of node A in directed graph G and C(A) is the number of outgoing links of A, define the PageRank formula. What is a typical value for the damping factor in the PageRank calculation?

(3 + 3) + 4 + 2 = 12

3. (a) Express Degree Centrality and Closeness Centrality with examples. What are the real world observations one may have about these two measures?

(b) Calculate Local and global clustering Coefficient for each node in the given network (you can name the nodes).



(4 + 2) + 6 = 12

Group – C

4. (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

5. (a) Derive the probability of observing exactly m edges in a random graph G(n,p).

(b) What is small-world effect? Derive and describe the equations for expressing degree distribution and clustering co-efficient for a small-world Network described by Watts-Strogatz.

(c) Critically comment on the statement and discuss: Hubs and power laws are absent in random networks.

(d) Find out the clustering co-efficient for a uniform circle graph (with n vertices), where, every node is connected to its c nearest neighbors (c/2 on each side).

3 + (1 + 3) + 2 + 3 = 12

Group – D

6. (a) Express advantages of SIR model over SI model. How outbreak can be defined in SIRS model?

(b) Illustrate the time evolution of SIR model with a diagram. How Epidemic can be identified by parameter values in SIR model?

(c) What is Influence? How is it affecting Social Networks?

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

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

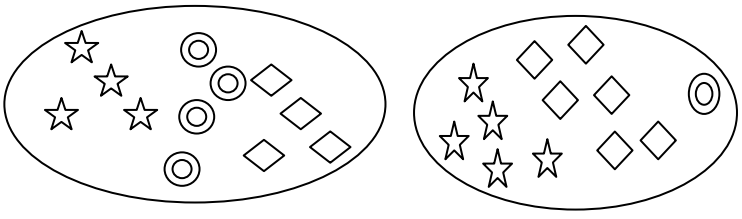
(b) Explain the nature of the infected compartment at time = t (I(t)) in terms of the total population (N) and the infected compartment at time = 0 (I(0)) along with the other model parameters in SI model (S should not be included). Comment on the nature of the curve.

4 + (7 + 1) = 12

Group – E

8. (a) What is a Web Community? Define Overlapping Community and hierarchy.

(b) Consider two clusters with 3 classes. Calculate NMI score.



(2 + 1 + 1) + 8 = 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.

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

Department & Section	Submission Link
CSE A + B + C	https://classroom.google.com/c/Mjk3NjI3 ODg3MDQ1/a/MzU5NDAyMDkzMTkx/details