LNMIIT

The LNM Institute of Information Technology
Department of Computer Science & Engineering
CSE 3132 Social Network Analysis
Exam Type: Final Exam

Time: 180 min 02/05/2019 **Max. Marks: 50**

Answer all questions in the same order as it appears in the question paper. Change in order may invite penalty! If there are any assumptions to be made for your answer write clearly the assumption that you are making before answering. Only if the assumption is reasonable it will be considered. **No doubt clarifications in the examination hall!** All the best!

1. Let $G(n,p)$ be a random graph which has a giant component in it and this was achieved with p being the least value. When we increase the number of nodes in $G(n,p)$ and you still want to have a giant component in it how the probability will vary? Will it increase, or decrease? Justify your answer very briefly. (3)

2. A cellular company ABC starts its cellular services in India. Their main objective is to reach a connected network – when we choose two people randomly from India they should be connected in the network through ABC. They keep track of average number of calls made or received by a person through ABC – say k. Their network does not distinguishes whether it is call made or received. Hence if they want to get a connected network with N people what should be the criteria? Justify your answer very briefly. (3)

   (a) $k > N/\log N$
   (b) $k < N/\log N$
   (c) $k > \log N$
   (d) $k < \log N$

3. Choose the correct answer(s) from the following statements based on a poison distribution for a random network and the distribution with respect to a scale-free network. Justify your choices briefly. (3)

   (a) The main difference between their distributions is in the tail of the degree distribution
   (b) The scale-free network has a large number of small degree nodes, most of which are absent in a random network
   (c) In a random network there are less number of nodes having degree around the average degree but that is not the case with a scale-free network
   (d) The probability of observing a high-degree node, or hub, is several orders of magnitude higher in a scale-free network than in a random network

4. What are the drawbacks of Barabasi-Albert (BA) Model? Choose the correct answer(s) from the following and justify the same. (3)

   (a) BA model captures triadic closures well
   (b) BA model does not explain the initial graph to start with
   (c) BA model works well with production of Scale-free graphs
   (d) BA model follows Preferential attachment technique and that captures small-world phenomena nicely

5. Suppose a company CDE would like to study the network of their customer support team (CST) that interacts regularly with the customers. Here the network has nodes that consists of their CST members and the customers who have interacted with the CST members over phone. Edges are formed between the nodes if they have interacted each other.

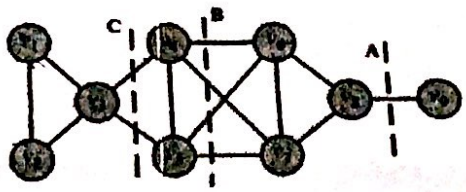
(a) For some reasons CDE wants to generate this network using an algorithm. Which algorithm is appropriate to use here? Why and How? (3)

(b) Now if the interaction between the CST and the customers was over a public forum will the same algorithm used to generate the network in the previous case will work? Justify your answer. (3)

6. (a) Give an example of a graph with local clustering coefficient 0 for all the vertices. Generalize your result. Justify your answer. (2)
(b) Give an example of a graph with local clustering coefficient 1 for all the vertices. Justify your answer. Generalize your result. Justify your answer. (2)
(c) What will be the local clustering coefficient of all the vertices in a wheel graph over n vertices. Generalize your result. Wheel graph over n vertices is defined as follows: (4)

- it is a cycle graph over $n-1$ vertices and
- the n^{th} node is adjacent to all other vertices.

7. (a) For the graph given below. Find the sizes of the cut at A, B and C. Which is the minimum cut size? Can this be used for the community detection algorithm as such? Why? (2)

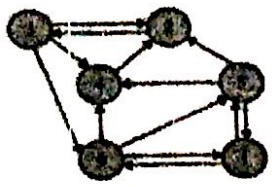


(b) For the same graph compute Jaccard and Cosine similarity between nodes v_4 and v_8, assuming that the neighborhood of a node excludes the node itself. (3)

(c) Compute Jaccard and Cosine similarity when the node is included in the neighborhood. (3)

8. Let G be a complete graph over $n = 4$ vertices. Using clique percolation method generate the clique graph for the value $k = 2$. Demonstrate the working of this clique percolation algorithm. Now try to generalize this. Given a complete graph K_n, how many nodes will the clique percolation method generate for the clique graph for value k? How many edges will it generate? (8)

9. (a) Follow the Independent Cascade Model procedure until it converges for the following graph. Assume that node i activates node j when $i - j \equiv 1(mod\ 3)$ and node 5 is activated at time 0. (4)



(b) Suppose one wants to study emotional contagion in a social network what kind of diffusion method will be appropriate to study it? Justify your answer with logical and mathematical argument, if any. (4)

All the best!