MA 653: Network Science
Spring 2025
Mathematics, IIT Guwahati

End-semester
Total Marks: 45
Duration: 180 mins

Answer all questions. Make suitable assumptions wherever necessary.
[All the symbols and variables used have usual meaning as discussed in class.]

A: Centrality Measures **[8 points]**

A.1 What is the key difference between degree centrality and eigenvector centrality? [2]

A.2 In which type of network: dense or sparse—is betweenness centrality generally higher, and why? [2]

A.3. What are the two main key issues with the traditional definition of closeness centrality? [2]

A.4. Consider a cycle (directed graph) of 3 nodes A → B → C → A, what is the eigenvector centrality of the nodes? Show your calculation. Assume the initial eigenvector centralities as $\mathbf{x} = [x_A, x_B, x_C]^T$. [2]

B: Configuration Model **[10 points]**

B.1. The configuration model can be thought of as the ensemble of all possible matchings of edge stubs, where node i has k_i stubs. Show that for a given degree sequence, the number of matchings is $(2m)!/(2^m m!)$. [3]

B.2. Suppose, we use the configuration model to build a friendship network. Show that the average degree of a neighbour (friend) is larger than the average degree of a node. [3]

B.3. In the configuration model, compute the probability of forming self-loops? [2]

B.4 Which structural property makes the configuration model a good approximation for real-world networks? [2]

C: Models of Network **[10 points]**

C.1. Consider a network created using the Price model, where the in-degree of vertex i is q_i. Let $p_q(n)$ be the fraction of vertices with in-degree q, when the total number of vertices in the network is n. Compute the value of p_0 (fraction of nodes with degree 0) in terms of c and a, after a new node is added to the network. [3]

C.2. Consider a modification of the Price model where c is 3. Each new node (paper) cites 2 old papers according to preferential attachment and 1 old paper uniformly at random. Would you still expect a power-law in-degree distribution? If yes, how might the exponent γ change compared to the actual preferential attachment case? If no, what will be the new exponent? [3]

C.3. What is the *first mover effect* in the general preferential attachment model? Explain intuitively why it happens. [2]

C.4. Consider the following variant of the Barabási-Albert model. Nodes are added one-by-one to a growing undirected network. Each node having initial degree c. The c edges emanating from a newly added node connect to previously existing nodes i with probability k_i+a. Given that c edges are added to the network with each node, what is the mean degree of a node in the network in the limit of large network size? [2]

D: Applications **10 points**

D.1. Consider the following modularity equation. Derive the equation for ΔQ, change in modularity, when we flip s_v to $-s_v$. [2]

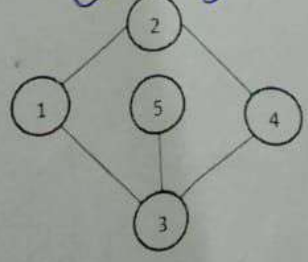
$$Q = \frac{1}{4m}\sum_{ij} B_{ij} s_i s_j \text{ where } B_{ij} = \left(A_{ij} - \frac{k_i k_j}{2m}\right)$$

D.2. When computing the change in modularity ΔQ, can we ignore terms like B_{vv} (i.e., the diagonal element) during flipping? Justify your answer. [2]

D.3. Suppose you are using the repeated bisection method for modularity maximization, and during the subdivision of a group, the modularity Q does not increase. Should you continue subdividing that group further? Why or why not? [2]

D.4. Consider a network as shown below, divided into two communities: Community A: {1, 2, 4} and Community B: {3, 5}. Consider a random walk 1, 2, 1, 3, 4, 3, 5. Calculate the entropy of the of the random walk, given by the map equation [4]

$$L = qH(Q) + \sum_g p_g H(P_g).$$



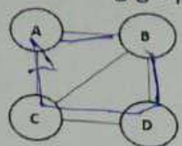
MA 653: Network Science
Spring 2025
Mathematics, IIT Guwahati

End-semester
Total Marks: 45
Duration: 180 mins

E: Machine Learning Meets AI

7 points

E.1. Consider the following graph [3]



The initial node embeddings are as follows: A: [1.0, 0.0], B: [0.5, 1.0], C: [1.0, 1.0] and D: [0.0, 1.0]. A single random walk of length 4 starting from node A is as follows: A→C→D→B. Using a context window of size 2, compute the negative log likelihood of node A.

$$L = \sum_{u \in V} \sum_{v \in N_R(u)} \left(-\log \frac{\exp(z_u^T z_v)}{\sum_{n \in V} z_u^T z_n} \right)$$

E.2. What are negative samples in shallow node embedding techniques? Why is negative sampling used instead of full softmax? [2]

E.3. A GNN layer computes: $h_v = ReLU(W.sum(\{h_u : u \in N(v)\}))$. Are the *sum* and *ReLU* function permutation invariant or permutation equivariant? Justify your answer briefly. [2]

---ooo---