SNA
G24AI1120
19 April, 2025

IITJ Major

G24AI1120
Abhinav Ghosh

Social Network Analysis

$E_1 \rightarrow A-B$
$E_2 \rightarrow B-C$
$E_3 \rightarrow B-D$

1) (a)

edge1	edge2	edge3
1	0	0
1	1	1
0	1	0
0	0	1

adjacency matrix

	A	B	C	D
A	0	1	0	0
B	1	0	1	1
C	0	1	0	0
D	0	1	0	0

(b) (B) Endos-Renyi (Random Network model)

It forms edges independently with uniform probability

(c) (c) Nash equilibrium

No player benefits from unilaterally changing their strategy.

(d) (B) Assortative mixing

It refers to preference for linking with similar objects.

(e) (D) Because it quantifies how often a node lies on the shortest path.

This helps identify critical path dependent node

(f) (c) The presence of many high degree hubs

Random failures miss hubs, targeted ones do not miss hubs.

(g) (A) The number of intra-community edges is significantly higher than expected in a random network with the same degree sequence.

high modularity seeks denser connections within groups.

(h) (B) 2/5

$|C \cap D| = 2$, $|union| = 5 \rightarrow$ Jaccard $= 2/5$

(i) (A) ICM uses edge probabilities independently, LTM uses a weighted sum of active neighbours compared to a node threshold.

LTM sums neighbour influence, ICM is probabilities per edge.

(j) (B) Because aggregating features from dissimilar neighbours can blur the node's own representative features, making classification harder.

This degrades classification in heterophilic settings

(2.) To minimize infection, ~~identity~~ we identify influential nodes ~~betwee~~ using betweeness centrality, which spots individuals on many shortest paths. These act as bridges in the network and are key in controlling spread. We then apply degree centrality to detect highly connected individuals who can spread quickly. Ranking nodes based on a hybrid of both centralities gives a more effective vaccination target set. Targetting only high degree nodes may ignore those bridging communities.

Betweeness helps protect against cross-cluster infection. With only 50% covered, this strategy maximises impact on SIR dynamics

(3.) We can use Node2Vec to generate vector embeddings capturing structural roles, and then we can combine with link prediction metrics like Jaccard ~~[illegible]~~ for scoring links.

Then train a model to learn from co-authorship & citation patterns.

Homophily drives most links - similar diad nodes collaborate often. To promote cross-disciplinary collaboration, we can add diversity-aware regularisation to scoring, like penalise recommending too many similar profiles, etc.

4. (a) Girvan-Newman algorithm detects communities of remove edges with highest betweeness, and this separates densely connected groups as bridges are eliminated. Iteratively it recalculates edge betweeness after each removal. and finally it also between high traffic link cut through also.

(b) It recalculats edge betweeness at each iteration Remove the edge with the highest value. to gradually separate communities.

(c) It is computationally expensive, like for large graphs & each step recalculates all pairwise shortest paths & this is a major disadvantage.

(d) The Louvain method optimises modularity through local moving of the nodes. It also then further aggregated community into super nodes. This process is fast & scalable for optimising modularity.

5. (a) Pagerank ranks nodes based on the importance of their neighbours, It distributes rank through links recursively. used widely in web search.

(f) The damping factor models random jump probability. It ensures that the surfer occasionaly jumps to a random node & it prevents trapping in loop/dead end. A higher damping factor gives more weight to link structure.

(c) Problem dangling nodes:-

Dangling nodes have no outbound links, and hence they break rank redistribution. This lead to rank sinks & prevents convergence.
Can be fixed by adding links to every node, which leads to convergence of the PageRank vector.

(6)

	strategy A	strategy B
Strategy U	(3,2)	(0,1)
Strategy L	(2,0)	(2,3)

(a) Best response of Player 1

If player 2 chooses A:-

U → 3 L → 2 → Best response is U

Player 2 chooses B:-

U → 0 L → 2 → Best response

Best response of player 2

player 1 choose U A → 2 B → 1 best response A

Player 1 choose L A → 0 B → 3 → Best response

(U,A), (L,B) are pure strategy Nash equilibrium.

PTO

(f) for player 1 playing
U with probability (p)
L with probability (1-p)

$\boxed{\text{given}}$

expected payoff :-

Strategy A
If P1 plays U : payoff = 2
P1 play L : payoff = 0
expected payoff = $2p + 0(1-p) = 2p$ ✓

Strategy B
P_1 play U : payoff = 1
P_1 play L : payoff = 3
Expected payoff : $1p + 3(1-p) = p - 3p + 3$
$= -2p + 3$

payoff

(c) using formulas :-

Strategy A : payoff = $2 \times 0.7 = 1.4$

✓ Strategy B : payoff :- $-2 \times 0.7 + 3 = 1.6$ ↑ higher

(7)

$$\begin{tikzcd} & B & \\ A \arrow[ur] & C \arrow[u] & D \arrow[ul] \end{tikzcd}$$

$N(B) = \{A, C, D\}$

given things in the question,

$$h_B^{(1)} = \sigma\left(W \cdot \left(\frac{1}{|N(B)|} \sum_{u \in N(B)} h_u^{(0)}\right)\right)$$

↳ this needs to be calculated.

σ is ReLU activation, weight matrix $W = \begin{pmatrix} 0.5 & 0 \\ 0.1 & 0.2 \end{pmatrix}$

initial vectors

$$h_A^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad h_C^{(0)} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \quad h_D^{(0)} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

① Aggregate neighbour features

$$\sum_{u \in N(B)} h_u^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

Average :-

$$\frac{1}{|N(B)|} \sum_{u \in N(B)} h_u^{(0)} = \frac{1}{3} \begin{bmatrix} 3 \\ 6 \end{bmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

② Transform by using linear transformation

$$W \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.5 \times 1 + 0 \times 2 \\ 0.1 \times 1 + 0.2 \times 2 \end{bmatrix} = \begin{pmatrix} 0.5 \\ 0.1 + 0.4 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$

③ Applying ReLU activation here:-

$$\mathrm{ReLU}\left(\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \right) = \begin{bmatrix} \max(0, 0.5) \\ \max(0, 0.5) \end{bmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$

$$h_B^{(1)} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$