CSCI 5352: Homework 2

Zachary Caterer[1,2,3]

[1]Department of Computer Science, University of Colorado Boulder

[2]Department of Chemical and Biological Engineering, University of Colorado Boulder

[3]Interdisciplinary Quantitative Biology Program, University of Colorado Boulder

February 8, 2025

Problem 1

Need to fix, complete, and check this proof.

Let the total number of nodes in the network be $n = n_A + n_B$. The closeness centrality of a node i is defined as:

$$C_i = \frac{n-1}{\sum_{j=1}^{n} d(i,j)}$$

where $d(i) = \sum_{j=1}^{n} d(i,j)$ represents the sum of shortest path distances from node i to all other nodes.

We aim to show the relationship between the closeness centralities C_A and C_B of vertices A and B.

Define Distance Sums The sum of distances from node A is:

$$d(A) = \sum_{j=1}^{n} d(A,j) = \sum_{j \in A} d(A,j) + \sum_{j \in B} d(A,j)$$

Similarly, for node B:

$$d(B) = \sum_{j=1}^{n} d(B,j) = \sum_{j \in B} d(B,j) + \sum_{j \in A} d(B,j)$$

Express Distance Relations Define:

$$S_A = \sum_{j \in A} d(A,j), \quad S_B = \sum_{j \in B} d(B,j)$$

where S_A is the total sum of distances from A to all nodes in A, and S_B is the sum of distances from B to all nodes in B.

Since the two subnetworks are connected by a single edge, each node in B is one unit farther from A than from B, and vice versa:

$$\sum_{j \in B} d(A,j) = \sum_{j \in B} (d(B,j) + 1) = S_B + n_B$$

$$\sum_{j\in A} d(B,j) = \sum_{j\in A}(d(A,j)+1) = S_A + n_A$$

Compute $d(A)$ and $d(B)$

$$d(A) = S_A + (S_B + n_B) = S_A + S_B + n_B$$

$$d(B) = S_B + (S_A + n_A) = S_A + S_B + n_A$$

Compute Closeness Centralities

$$C_A = \frac{n-1}{d(A)} = \frac{n-1}{S_A + S_B + n_B}$$

$$C_B = \frac{n-1}{d(B)} = \frac{n-1}{S_A + S_B + n_A}$$

Taking reciprocals:

$$\frac{1}{C_A} = \frac{S_A + S_B + n_B}{n-1}$$

$$\frac{1}{C_B} = \frac{S_A + S_B + n_A}{n-1}$$

Dividing both equations:

$$\frac{\frac{1}{C_A}}{\frac{1}{C_B}} = \frac{S_A + S_B + n_B}{S_A + S_B + n_A} = \frac{n_B}{n_A}$$

Rearranging:

$$\frac{1}{C_A} n_A = \frac{1}{C_B} n_B$$

Multiplying by $\frac{1}{n}$:

$$\frac{1}{C_A} n_A \frac{1}{n} = \frac{1}{C_B} n_B \frac{1}{n}$$

Thus, we have proven the required relationship:

$$\frac{1}{C_A}\frac{n_A}{n} = \frac{1}{C_B}\frac{n_B}{n}$$

Problem 2

Need to fix, complete, and check this section.

Part A

1. Selection Criteria:
   - Randomly select two directed edges (u, v) and (x, y) from the network.
   - Since it is a directed network, the order of the nodes in the edge matters. Thus, the two edges are distinct.
2. Possible output configurations:
   - Option 1: $(u, x)$and (v, y)
   - Option 2: $(u, y)$and (v, x)
3. Checkes needed on G':
   - Ensure no self loops.
   - Ensure no multiple edges.
   - Verify the k^{in} and k^{out} values of the nodes remain unchanged.

Part B

1. Selection Criteria:
   - Randomly select two directed edges one from each partition, for simplicity sake lets say nodes u and v are from partition A and nodes x and y are from partition B. So the two edges are (u, x) and (v, y).
2. Possible output configurations:
   - The only option: (u, y) and (v, x)
3. Checkes needed on G':
   - Ensure no self loops.
   - Ensure no multiple edges.
   - Verify the degrees of the nodes remain unchanged.
   - Veriy the new edges maintain bipartite structure.

Problem 3

Need to fix, complete, and check this section.

Problem 4

Problem 5

Part A

Part B

Part C

Problem 6

Problem 7

- What paper did you choose?
- What was the research question?
- What was the approach the authors took to answer that question?
- What did they do well?
- What could they have done better?
- What extensions can you envision?