Introduction to Economic and Social Networks, Problem Set I

Chen Gao*

October 11, 2025

1 Q1

From the graph, it's easy to get the shortest path distance matrix

$$L = \left(\ell_{ij}\right)_{6\times 6} = \begin{bmatrix} 0 & 1 & 1 & 2 & 3 & 3 \\ 1 & 0 & 1 & 2 & 3 & 3 \\ 1 & 1 & 0 & 1 & 2 & 2 \\ 2 & 2 & 1 & 0 & 1 & 1 \\ 3 & 3 & 2 & 1 & 0 & 2 \\ 3 & 3 & 2 & 1 & 2 & 0 \end{bmatrix}$$

For node 3

1. closeness centrality is given by

$$C(3) = \frac{n-1}{\sum_{j\neq 3} \ell_{3j}} = \frac{5}{1+1+1+2+2} = \frac{5}{7}$$

*National School of Development, Peking University. Email chengao0716@gmail.com

2. decay centrality (with $\delta = 0.5$) is given by

$$
\begin{aligned}
C_3^{\delta}(g) &= \frac{1}{(n-1)\,\delta} \sum_{j \neq 3} \delta^{\,\ell(3,j)} \\
&= \frac{1}{(n-1)\,\delta} \left(3\delta + 2\delta^2\right) \\
&= \frac{2}{5}\,(1.5 + 2 \times 0.25) \\
&= \frac{4}{5}
\end{aligned}
$$

3. degree centrality is given by

$$
C_D\,(3) = \frac{3}{n-1} = \frac{3}{5}
$$

4. betweenness centrality

$$
C_3^B(g) = \frac{2}{(n-1)\,(n-2)} \sum_{\substack{j<k \\ j,k \neq 3}} \frac{P_3(j,k)}{P_{j,k}}
$$

Node 3 lies on the shortest path for the following pairs: (1, 4), (1, 5), (1, 6), (2, 4), (2, 5), (2, 6). For each of these 6 pairs, there is only one shortest path, and it passes through node 3. So

$$
C_3^B(g) = \frac{2}{(n-1)\,(n-2)} \sum_{\substack{j<k \\ j,k \neq 3}} \frac{P_3(j,k)}{P_{j,k}} = \frac{1}{10} \times 6 = 0.6
$$

For node 5

1. closeness centrality is given by

$$
C(5) = \frac{n-1}{\sum_{j \neq 5} \ell_{5j}} = \frac{5}{3+3+2+1+2} = \frac{5}{11}
$$

2. decay centrality (with $\delta = 0.5$) is given by

$$\begin{aligned} C_5^{\delta}(g) &= \frac{1}{(n-1)\,\delta} \sum_{j \neq 5} \delta^{\,\ell(5,j)} \\ &= \frac{1}{(n-1)\,\delta} \left(\delta + 2\delta^2 + 2\delta^3\right) \\ &= \frac{2}{5}\,(0.5 + 0.5 + 0.25) \\ &= 0.5 \end{aligned}$$

3. degree centrality is given by

$$C_D\,(5) = \frac{1}{5}$$

4. betweenness centrality: Node 5 is a leaf node (it only connects to one other node). Therefore, it cannot lie on the shortest path between any two other nodes.

$$C_5^B(g) = \frac{2}{(n-1)\,(n-2)} \sum_{\substack{j<k \\ j,k \neq 5}} \frac{P_5(j,k)}{P_{j,k}} = 0$$

Calculate the network's average path length Given L, we have:

$$\begin{aligned} L &= \frac{1}{\frac{N(N-1)}{2}} \sum_{i<j} \ell(i,j) \\ &= \frac{1}{3 \times 5} \left[1+1+2+3+3+1+2+3+3+1+2+2+1+1+2\right] \\ &= \frac{28}{15} \end{aligned}$$

2 Q2

Node 1's degree Given the graph, we have

$$k_1 = 8$$

Node 1's clustering coefficient The maximum number of edges that can exist between neighbors is $\frac{k_1(k_1-1)}{2} = \frac{8\times7}{2} = 28$, and the 4 neighbors on the left themselves form a complete graph, having $C_4^2 = 6$ edges. There is no connection between the 4 leaf node neighbors on the right. There is also no direct connection between the left neighbor and the right neighbor. So $E_1 = 6$ and

$$C_1 = \frac{2E_1}{k_1(k_1-1)} = \frac{6}{28} = \frac{3}{14}$$

Node 2's clustering coefficient Node 2 is in the 5-node complete graph on the left. Its neighbors are the other 4 nodes in the complete graph (including node 1). We have

$$k_2 = 4$$

Since these four neighbors are in the same complete graph, they are connected to each other. So $E_2 = C_4^2 = 6$,

$$C_2 = \frac{2E_2}{k_2(k_2-1)} = \frac{6}{6} = 1$$

The network's (including all 9 nodes) overall clustering All triangles are in the K5 complete graph on the left. The number of triangles in a K5 graph is

$$\binom{5}{3} = \frac{5\times4\times3}{3\times2\times1} = 10$$

Note the triples centered on the four leaf nodes on the right: their degrees are all 1 and cannot form a connected triple.

of the triplet centered at node 1:

$$\binom{k_1}{2} = \binom{8}{2} = \frac{8 \times 7}{2} = 28$$

of the triplet centered on the remaining 4 nodes on the left (including node 2)

$$4 \times \binom{4}{2} = 24$$

so the network's (including all 9 nodes) overall clustering is

$$C = \frac{3 \times 10}{28 + 24} = \frac{30}{52} = \frac{15}{26}$$

The network's average clustering The coefficients of the four nodes in K5 on the left, except node 1, are all 1. The degree of the four leaf nodes on the right is 1. According to the requirements of the question, their clustering coefficient is considered to be 0. So

$$C_{avg} = \frac{1}{N}\sum_{i=1}^{N} C_i = \frac{1}{9}\left(\frac{3}{14} + 4 \times 1 + 4 \times 0\right) = \frac{1}{9}\left(\frac{3}{14} + 4\right) = \frac{1}{9}\left(\frac{59}{14}\right) = \frac{59}{126}$$

As $k \to +\infty$, calculate the network's overall clustering In this case, on the left is a $k+1$ complete graph K_{k+1}, node 1 is connected to k leaf nodes, so:

# of Triangular is

$$\binom{k+1}{3} = \frac{(k+1)k(k-1)}{6}$$

# of connected triplets:

$$\binom{2k}{2} + k \times \binom{k}{2} = k\left(2k-1\right) + \frac{k^2\left(k-1\right)}{2}$$

so

$$
\begin{aligned}
C(k) &= \frac{3 \times \binom{k+1}{3}}{\binom{2k}{2} + k \times \binom{k}{2}} = \frac{3 \times \frac{(k+1)k(k-1)}{6}}{k(2k-1) + \frac{k^2(k-1)}{2}} \\
&= \frac{(k+1)k(k-1)}{2k(2k-1) + k^2(k-1)}
\end{aligned}
$$

and

$$
\lim_{k \to \infty} C(k) = \lim_{k \to \infty} \frac{(k+1)k(k-1)}{2k(2k-1) + k^2(k-1)} = 1
$$

As $k \to +\infty$, calculate the network's average clustering Given

$$
C_{avg}(k) = \frac{1}{N} \sum_{i=1}^{2k+1} C_i,
$$

we have $C_i = 0$ for k leaf nodes . And the k nodes in K_{k+1} except node 1: C_i=1.For node 1, we have

$$
C_1 = \frac{\binom{k}{2}}{\binom{2k}{2}} = \frac{k(k-1)/2}{k(2k-1)} = \frac{k-1}{2(2k-1)}
$$

so

$$
\begin{aligned}
C_{avg}(k) &= \frac{1}{2k+1} \left[\frac{k-1}{2(2k-1)} + k \right] \\
&= \frac{1}{2k+1} \frac{k-1+k(4k-2)}{2(2k-1)} \\
&= \frac{1}{2k+1} \frac{4k^2-k-1}{2(2k-1)} \\
&= \frac{4k^2-k-1}{2(4k^2-1)}
\end{aligned}
$$

so

$$
\lim_{k \to \infty} C_{avg}(k) = \lim_{k \to \infty} \frac{4k^2-k-1}{2(4k^2-1)} = \frac{1}{2}
$$

3 Q3

The adjacency matrix By definition, we have

$$G = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

The eigenvector centrality for each node We solve the characteristic equation

$$\det(G - \lambda I) = 0$$

which is

$$\begin{aligned} \det(G - \lambda I) &= (-\lambda) \cdot C_{11} + (1) \cdot C_{12} + (1) \cdot C_{13} + (1) \cdot C_{14} \\ &= \lambda^4 - \lambda^2 - \lambda^2 - \lambda^2 \\ &= \lambda^4 - 3\lambda^2 = 0 \end{aligned}$$

The eigenvalues for this star network are $\sqrt{3}, -\sqrt{3}, 0, 0$. So the spectral radius is

$$\rho\left(G\right) = \sqrt{3}$$

we try to solve

$$Gx = \rho\left(G\right)x$$

that is

$$\begin{aligned} x_2 + x_3 + x_4 &= \sqrt{3}x_1 \\ x_1 &= \sqrt{3}x_2 \\ x_1 &= \sqrt{3}x_3 \\ x_1 &= \sqrt{3}x_4 \end{aligned}$$

it's easy to see that x is proportional to $(\sqrt{3},1,1,1)'$, taking L_2norm for normalization, we have

$$\mathbf{c} = [c_1, c_2, c_3, c_4]' = \frac{1}{\sqrt{6}}\left[\sqrt{3},1,1,1\right]'$$

The Katz-Bonacich centrality of each node, when $b = 0.6$ and $a = 1/2$ The Katz-Bonacich centrality is calculated using the formula

$$C_{KB} = a(I - bG)^{-1}\mathbf{1}$$

Since

$$\begin{aligned} I - bG &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - b\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -0.6 & -0.6 & -0.6 \\ -0.6 & 1 & 0 & 0 \\ -0.6 & 0 & 1 & 0 \\ -0.6 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

given

$$\begin{aligned} C_{KB} &= a(I - bG)^{-1}\mathbf{1} \\ (I - bG)\, C_{KB} &= a\mathbf{1} \end{aligned}$$

and let $C_{KB} = [C_1, C_2, C_3, C_4]'$, we have

$$\begin{aligned} C_1 - 0.6C_2 - 0.6C_3 - 0.6C_4 &= \tfrac{1}{2} \\ -0.6C_1 + C_2 &= \tfrac{1}{2} \\ -0.6C_1 + C_3 &= \tfrac{1}{2} \\ -0.6C_1 + C_4 &= \tfrac{1}{2} \end{aligned}$$

which is solved by

$$C_{KB} = [C_1, C_2, C_3, C_4]' = [-17.5, -10, -10, -10]'$$

Counting walks It's easy to see that

$$G^2 = G \times G = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}$$

and

$$G^3 = G^2 \times G = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 3 & 3 \\ 3 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{pmatrix}$$

so we have

1. The number of walks of length 2, from node 2 to 3 is $G^2_{23} = 1$
2. The number of walks of length 3, from node 1 to 4 is $G^3_{14} = 3$

The diffusion centrality for each node, when $b = 0.5$ and $T = 3$ By definition, we have

$$C_D = \sum_{t=1}^{3} (0.5G)^t \mathbf{1} = ((0.5G)^1 + (0.5G)^2 + (0.5G)^3)\mathbf{1}$$

so we have

$$0.5G + 0.25G^2 + 0.125G^3 = \begin{pmatrix} 0.75 & 0.875 & 0.875 & 0.875 \\ 0.875 & 0.25 & 0.25 & 0.25 \\ 0.875 & 0.25 & 0.25 & 0.25 \\ 0.875 & 0.25 & 0.25 & 0.25 \end{pmatrix}$$

so we have

$$C_D = [3.375, 1.625, 1.625, 1.625]'$$

4 Q4

QA A diameter of 2 means the radius $r = 1$. This describes a star net. Given each non-leaf node has a degree $d = 100$, we have $n = 1 + 100 = 101$

QB We can calculate the total number of nodes by summing the nodes at each "level" away from the center.

It easy to see that the number of nodes at level $k = 1, 2, ..., r$ is $d \times (d-1)^{k-1}$, so

$$\begin{aligned}
n &= 1 + \sum_{k=1}^{r} d \times (d-1)^{k-1} \\
&= 1 + d \sum_{k=1}^{r} \times (d-1)^{k-1} \\
&= 1 + d \times \frac{(d-1)^r - 1}{(d-1) - 1} \\
&= 1 + d \frac{(d-1)^r - 1}{d-2} \\
&= \frac{d(d-1)^r - 2}{d-2} = \frac{100 \times 99^r - 2}{98}
\end{aligned}$$

QC Given $n = \frac{d(d-1)^r - 2}{d-2}$ we have

$$\begin{aligned}
n &= \frac{d(d-1)^r - 2}{d-2} \\
&\approx \frac{d(d-1)^r}{d-2} \\
&\approx (d-1)^r
\end{aligned}$$

so we have $\ln n \approx r \ln(d-1)$ or equivalently [1]

$$D \approx 2\frac{\ln n}{\ln(d-1)}$$

QD Given the result in C, we have

$$\begin{aligned}
D &\approx 2\frac{\ln 5 \times 10^9}{\ln 99} \\
&\approx 2\log_{99} 5 \times 10^9 \\
&\approx 2\log_{100} 5 \times 10^9 \\
&\approx \log_{10} 5 \times 10^9 \\
&\approx \log_{10} 5 + 9 \\
&\approx 0.7 + 9 \approx 10
\end{aligned}$$

Note that a diameter of 9 would not be enough to contain 5 billion nodes, so we must round up to the next integer.

[1] In the problem statement we have "provide an approximation of the diameter, r." To avoid notation conflict, I use D here to denote the diameter.