Data Analytics for Graphs COMP9312 22T2

## **Answer for Assignment 2**

## Q1:

1. The graphs in Figure 1 and Figure 2 are isomorphic. They have the following isomorphism to map graph in figure 1 to figure 2:

Figure 1	Figure2
1	1
2	2
3	6
4	4
5	3
6	7
7	8
8	5

2. **Node A** does not lie in any shortest path other than itself. A has shortest path: A-B, A-C, A-B-D, A-B-D-F, A-B-D-E, A-B-D-F, A-B-D-G-H

Betweenness centrality of node A is 0

Closeness centrality of node A is  $\frac{1}{21}$ 

**Node D** lies in the shortest path of the following: (1) B to all the other nodes on the right of D has to pass through D: 5. (2). C to all the other nodes on the right of D has to pass through D: 5. (3). A to all the other nodes on the right of D has to pass through D, which has B or C in the path: 5. (4). F to E has two shortest paths:  $F \to D \to E$ ,  $F \to G \to E$ : 0.5.

D has shortest paths: D-B-A, D-B, D-C, D-E, D-F, D-G, D-F-I, D-G-H.

Betweenness centrality of node D is 5 + 5 + 5 + 0.5 = 15.5

Closeness centrality of node D is  $\frac{1}{11}$ 

**Node G** lies in the shortest path of the following: (1) H to all the other nodes has to pass through G: 7. (2). A to I: 0.5, B to I: 0.5, C to I:0.5, D to I:0.5, E to I: 1. (3). F to E has two shortest path:  $F \to D \to E$ ,  $F \to G \to E$ : 0.5.

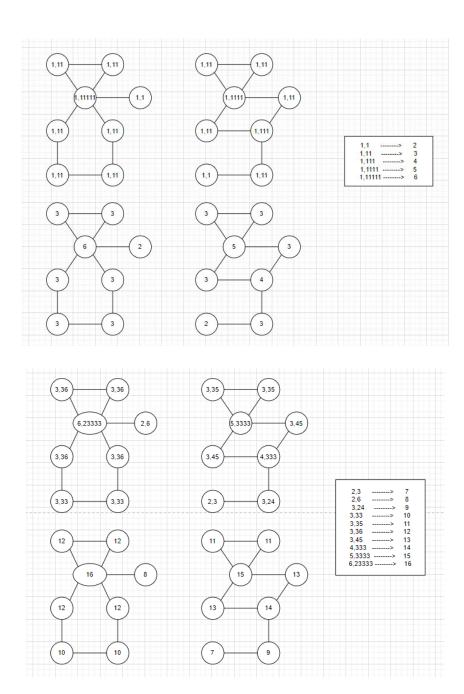
-

G has the shortest paths: G-I, G-H, G-E, G-F, G-D, G-D-B, G-D-C, G-D-C-A.

Betweenness centrality of node G is 7 + 3 + 0.5 = 10.5

Closeness centrality of node G is  $\frac{1}{12}$ 

3.



Data Analytics for Graphs COMP9312 22T2

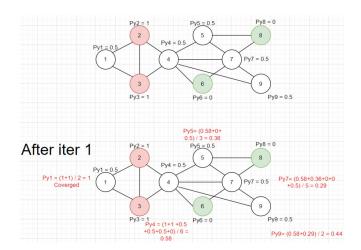
$$\mathbf{\varphi(G1)} = \begin{bmatrix} 1,2,3 & 4,5,6,7,8,9,10,11,12,13,14,15,16 \\ \mathbf{\varphi(G1)} &= \begin{bmatrix} 8,1,6,0,0,1,0,0,2,0,4,0,0,0,1,1 \end{bmatrix} \\ 1,2,3 & 4,5,6,7,8,9,10,11,12,13,14,15,16 \\ \mathbf{\varphi(G2)} &= \begin{bmatrix} 8,1,5,1,1,0,1,1,0,2,0,2,1,1,0,0 \end{bmatrix} \end{bmatrix}$$

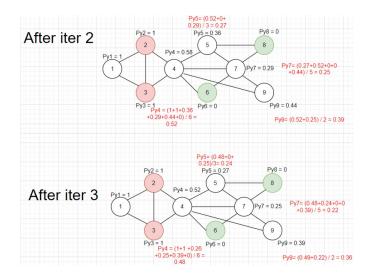
$$K(G1,G2) = \phi(G1)^T \phi(G2) = 95.$$

## **Q2**:

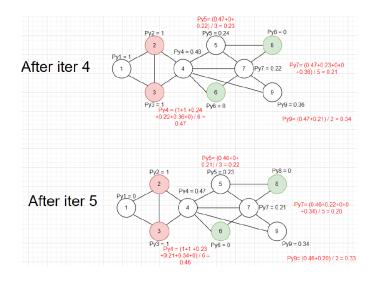
1.

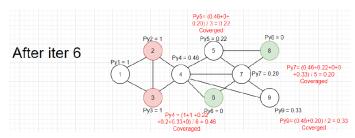
It can reach the convergence, 6 rounds it should take, I will update the label from smallest to largest (from 1 to 9 without those converged nodes).





Data Analytics for Graphs COMP9312\_22T2





Node 1 will be red (> 0.5 and has label 1) and node 3, 5, 7, 9 will be green (< 0.5 and has label 0).

2.

Target t	Probability	Dist(3,t)
3	1/4	0
2	1	1
5	1/3	2
6	1/3	2
7	1/3	2
9	1/3	2

Q3:

The result is:

$$\begin{bmatrix} 0.5 & 1.3 & 0.7 & 0.9 \\ 1.33 & 1.27 & 1.5 & 0.53 \\ 1.87 & 0.067 & 0.97 & 0.87 \\ 1.5 & 0.35 & 1.3 & 0.65 \\ 1.63 & 0.23 & 1.43 & 0.1 \\ 1 & 1.3 & 1.4 & 0.7 \\ 0.82 & 0.94 & 1.46 & 0.18 \\ 0.95 & 0.15 & 0.9 & 0.2 \\ 1 & 1 & 2 & 0 \\ \end{bmatrix}$$

Q4:

1.

$$cos(u_4, u_5) = \frac{1.1223}{1.1695} = 0.9596$$
$$cos(u_2, u_7) = \frac{0.7175}{1.0498} = 0.6835$$

2. If t=0.6, node pairs (u4, u5) and (u2, u7) will be predicted as there is a link, however, in the graph above, there is no link between u2 and u7, so the prediction for (u2, u7) is wrong. The prediction for (u4, u5) is correct, there has a link between u4 and u5 in the graph. If the threshold become t=0.8, it will predict that there has a link between u4 and u5, but there has no link between u2 and u7. This is correct for the graph.