Written Assignment 4 solution
SENG 474/CSC 578D

April 12, 2019

Question 1

a.

$Similarity(U_1, U_2) = \dfrac{|U_1 \cap U_2|}{|U_1 \cup U_2|}$

Following up with the formula, the similarity matrix will be,

$$\begin{bmatrix} 1 & 0 & 2/3 & 1/3 \\ 0 & 1 & 1/4 & 1/3 \\ 2/3 & 1/4 & 1 & 2/3 \\ 1/3 & 1/3 & 2/3 & 1 \end{bmatrix}$$

b.

$Pref(A, I_2) = S[A, B] + S[A, C] + S[A, D] = 0 + \dfrac{2}{3} + \dfrac{1}{3} = 1$

$Pref(A, I_3) = S[A, B] = 0$

$Pref(D, I_1) = S[A, D] + S[C, D] = \dfrac{1}{3} + \dfrac{2}{3} = 1$

$Pref(D, I_2) = S[B, D] = \dfrac{1}{3}$

Question 2

a.

$$U_0V_0 = \begin{bmatrix} 1 & 4 & 6 & 2 \\ 2 & 5 & 6 & 1 \\ 1 & 3 & 4 & 1 \\ 2 & 4 & 4 & 0 \end{bmatrix} \quad M = \begin{bmatrix} 2 & & 4 & 1 \\ & 1 & 3 & \\ 2 & 4 & & 3 \\ 1 & 5 & & \end{bmatrix}$$

Considering only the non empty positions in M, and adding up the squared difference ,

$$RMSE = \sqrt{\frac{(1-2)^2+(6-4)^2+(2-1)^2+(5-1)^2+(6-3)^2+(1-2)^2+(3-4)^2+(1-3)^2+(2-1)^2+(4-5)^2}{10}} = \sqrt{\frac{39}{10}}$$

b.

$$UV = \begin{bmatrix} x_1 & 2x_1+2 & 2x_1+4 & 2 \\ x_2 & 2x_2+1 & 2x_2+2 & 1 \\ x_3 & 2x_3+1 & 2x_3+2 & 1 \\ x_4 & 2x_4 & 2x_4 & 0 \end{bmatrix} M = \begin{bmatrix} 2 & & 4 & 1 \\ & 1 & 3 & \\ 2 & 4 & & 3 \\ 1 & 5 & & \end{bmatrix}$$

$$RMSE = \sqrt{\frac{(x_1-2)^2+(2x_1)^2+(1)^2+(2x_2)^2+(2x_2-1)^2+(x_3-2)^2+(2x_3-3)^2+(-2)^2+(x_4-1)^2+(2x_4-5)^2}{10}}$$

Let $y = (x_1-2)^2+(2x_1)^2+(1)^2+(2x_2)^2+(2x_2-1)^2+(x_3-2)^2+(2x_3-3)^2+(-2)^2+(x_4-1)^2+(2x_4-5)^2$

In order to minimize RMSE, y should be minimized

$=> \frac{\partial(y)}{\partial(x_1)} = 0, \frac{\partial(y)}{\partial(x_2)} = 0, \frac{\partial(y)}{\partial(x_3)} = 0, \frac{\partial(y)}{\partial(x_4)} = 0$

Upon solving these equations, $x_1 = \frac{2}{5}, x_2 = \frac{1}{4}, x_3 = \frac{8}{5}, x_4 = \frac{11}{5}$

c.

$$U = \begin{bmatrix} 2/5 & 2 \\ 1/4 & 1 \\ 8/5 & 1 \\ 11/5 & 0 \end{bmatrix} V = \begin{bmatrix} 1 & 2 & 2 & 0 \\ y_1 & y_2 & y_3 & y_4 \end{bmatrix} UV = \begin{bmatrix} 2/5+2y_1 & 4/5+2y_2 & 4/5+2y_3 & 2y_4 \\ 1/4+y_1 & 1/2+y_2 & 1/2+y_3 & y_4 \\ 8/5+y_1 & 16/5+y_2 & 16/5+y_3 & y_4 \\ 11/5 & 22/5 & 22/5 & 0 \end{bmatrix}$$

$$RMSE = \sqrt{\frac{(2y_1-\frac{8}{5})^2+(y_1-\frac{2}{5})^2+\frac{6}{5})^2+(y_2-\frac{1}{2})^2+(y_2-\frac{4}{5})^2+(\frac{-3}{5})^2+(2y_3-\frac{16}{5})^2+(y_3-\frac{5}{2})^2+(2y_4-1)^2+(y_4-3)^2}{10}}$$

Let $z = (2y_1-\frac{8}{5})^2+(y_1-\frac{2}{5})^2+\frac{6}{5})^2+(y_2-\frac{1}{2})^2+(y_2-\frac{4}{5})^2+(\frac{-3}{5})^2+(2y_3-\frac{16}{5})^2+(y_3-\frac{5}{2})^2+(2y_4-1)^2+(y_4-3)^2$

In order to minimize RMSE, z should be minimized

$=> \frac{\partial(z)}{\partial(dy_1)} = 0, \frac{\partial(z)}{\partial(y_2)} = 0, \frac{\partial(z)}{\partial(y_3)} = 0, \frac{\partial(z)}{\partial(y_4)} = 0$

Upon solving these equations, $y_1 = \frac{36}{50}, y_2 = \frac{13}{20}, y_3 = \frac{89}{50}, y_4 = 1$

Question 3

a.

If we consider shortest paths only from A, then there is only one shortest path through the edges HI, HG, DE and DF. Edges BH and CD contribute to 3 shortest path each. BH lies on the shortest path from A to I, A to G and A to H. CD lies on the shortest path from A to D, A to E and A to F. Edges AC and AB contribute to 4 shortest path each. AB lies on the shortest path from A to B,H,I,G and AC lies on the shortest path from A to C,D,E,F

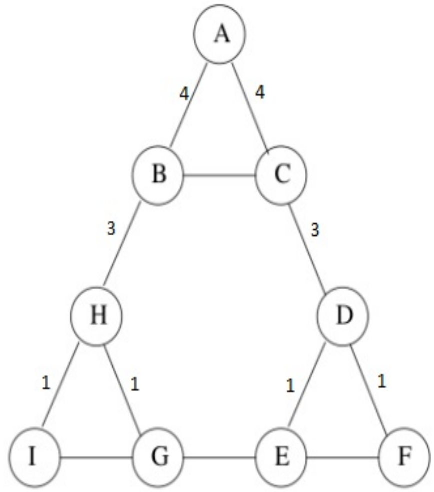

Figure 1: part a

b.

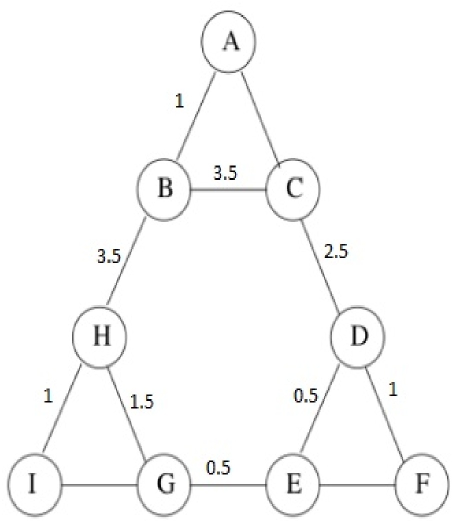

Figure 2: part b

Since there are two shortest paths from B to E, B-C-D-E and B-H-G-E , the edges GE and DE both get half the contribution. Following up with the girvan newman algorithm, the result will be similar to the shown in figure 2.

c.

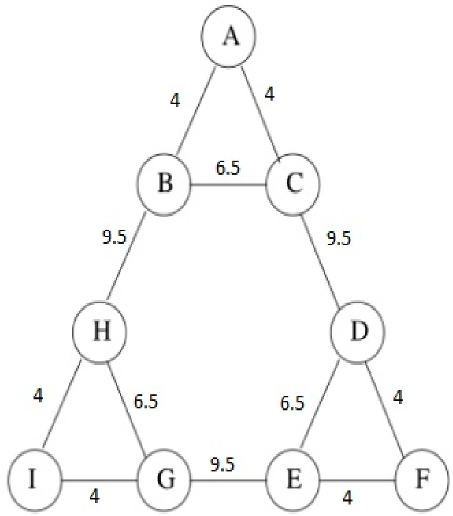

Figure 3: part c

Considering the symmetry, it can be noticed that the following pair of edges will have the same betweenness:-
AB, AC, HI, IG, EF, DF
BC, DE, HG
BH, GE, CD
So, let's pick only edge one from each set and calculate it's betweenness by adding up contributions by the shortest paths from different nodes

$Sourcenode$	(AB)	(BC)	(BH)
A	4	0	3
B	1	3.5	3.5
C	0	3.5	2.5
D	0	1.5	0.5
E	0	0.5	0.5
F	0	1	0
G	1	0.5	2.5
H	1	1.5	3.5
I	1	1	3

$$betweenness(AB) = \frac{4+1+0+0+0+0+1+1+1}{2} = 4$$
$$betweenness(BC) = \frac{0+3.5+3.5+1.5+0.5+1+0.5+1.5+1}{2} = 6.5$$
$$betweenness(BH) = \frac{3+3.5+2.5+0.5+0.5+0+2.5+3.5+3}{2} = 9.5$$

Question 4

Number of possible edges with 5 nodes = $\binom{5}{2} = 10$
Number of actual edges present = 8
If the probability of presence of an edge is p, then the probability of observing the same graph as in the figure, $f(p) = p^8(1-p)^2$
Maxima of f(p) would follow, $\frac{\partial f(p)}{\partial (p)} = 0$
$\frac{\partial f(p)}{\partial (p)} = 8p^7(1-p)^2 - 2p^8(1-p) = 2p^7(1-p)(4-4p-p) = 2p^7(1-p)(4-5p)$
If f(p) = 0 , then the possible values of p are 0,1 and $\frac{4}{5}$
But when p=0 or p=1,then f(p)=0 and represents the least likelihood
Thus p=0, p=1 actually minimize f(p) and the only value that maximizes it is $p = \frac{4}{5}$

Question 5

By observing the graph, it can be noticed that there are 9 triangles in total:-

$U_1 \to T_1 \to W_1$
$U_1 \to T_2 \to W_1$
$U_1 \to T_2 \to W_2$
$U_1 \to T_3 \to W_2$
$U_2 \to T_2 \to W_2$
$U_2 \to T_2 \to W_3$
$U_2 \to T_3 \to W_2$
$U_2 \to T_4 \to W_2$
$U_2 \to T_4 \to W_3$

Question 6

$\mathbf{M} = \begin{bmatrix} 1 & 1 \\ 8 & 3 \end{bmatrix}$

If λ is an eigen value of $\mathbf{M}$, then it should satisfy $|\mathbf{M} - \lambda\mathbf{I}| = 0$, where $\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\mathbf{M} - \lambda\mathbf{I} = \begin{bmatrix} 1-\lambda & 1 \\ 8 & 3-\lambda \end{bmatrix}$
$|\mathbf{M} - \lambda\mathbf{I}| = (1-\lambda)(3-\lambda) - 8 = \lambda^2 - 4\lambda - 5$
$\implies (\lambda+1)(\lambda-5) = 0$
$\implies \lambda_1 = -1, \lambda_2 = 5$

If $\mathbf{e}$ is an eigen vector of $\mathbf{M}$, it should satify $(\mathbf{M} - \lambda\mathbf{I})\mathbf{e} = \mathbf{0}$, where $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Let $\mathbf{e} = \begin{bmatrix} x \\ y \end{bmatrix}$

$Then,$ $\begin{bmatrix} 1-\lambda & 1 \\ 8 & 3-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \begin{bmatrix} (1-\lambda)x + y \\ 8x + (3-\lambda)y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$(1-\lambda)x + y = 0 \tag{1}$$

$$8x + (3-\lambda)y = 0 \tag{2}$$

If λ=5, then y= 4x
If λ=-1, then y=-2x

Thus the eigen vector corresponding to eigen value 5 , will be of the form $\begin{bmatrix} x \\ 4x \end{bmatrix}$

And the eigen vector corresponding to the eigen value -1 will be of the form $\begin{bmatrix} x \\ -2x \end{bmatrix}$

Matrices can have more than one eigenvector sharing the same eigenvalue and this example shows it explicitly.