

### Assignment (Part one Week 3)

#### Question 1

Use the whatapps, or Facebook to explain the 4V concept in BigData.

(20 marks)

#### Velocity:

Velocity means the speed of data generated is extremely fast.

In each second, people from all around the world keep uploading posts and videos to Facebook. When you keep refreshing, each refresh will give you a new post and video. It is also doesn't stop generate when we are sleeping.

#### Volume:

Volume means the file size and the amount of data generated each day are usually significant.

In Facebook, 500+ terabytes of data are being generated each day. It's pulling in 2.7 billion Like actions and 300 million photos per day, and it scans roughly 105 terabytes of data each half-hour. This lager data is hard to process by laptop and desktop.

#### Variety:

Variety means the data comes from a different source in different kinds of form and not only created by humans but also the machine.

On Facebook, there are different kinds of data, like text, video, audio, and images. People are creating them in a different place. Also, the server of Facebook creates a server log and user log. All of these data are being sorted and stored on the Facebook server.

#### Veracity:

Veracity means the quality of the data; they are useful to us or not. High veracity data are being valuable for us to analyze and give a meaningful result to us. Low veracity data usually are meaningless data.

When there were accidents happened, some people usually open their phones and start streaming. The live video is useful for police to investigate the accident. Also, the users' search record can let Facebook calculate user's hobbies and provide more that fits the user's hobby recommend post. On the other side, some posts on Facebook are about user's own life, and some may be deleted by users. Those data cant being analyzed by Facebook and being useless.

Question 2

EXE 1 Question 6

(5 marks)

$$6 \times I1 + 3(I1 - I2) = 42$$

$$6I1 + 3I1 - 3I2 = 42$$

$$9I1 - 3I2 - 42 = 0 \text{ ---(1)}$$

$$4 \times I2 + 3(I2 - I1) = 10$$

$$4I2 + 3I2 - 3I1 = 10$$

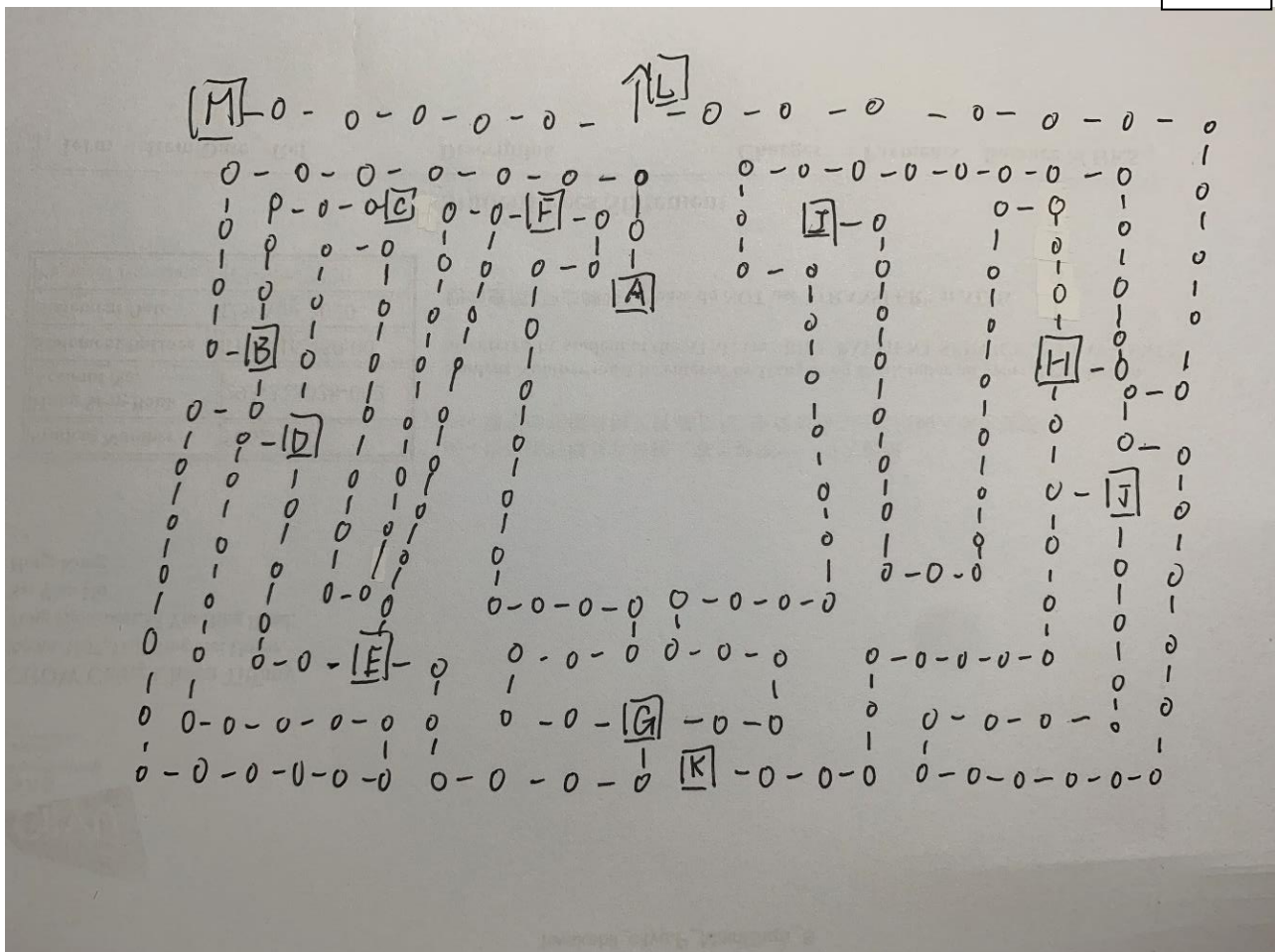
$$-3I1 + 7I2 - 10 = 0 \text{ ---(2)}$$

By solving the equation,

$$I1 = 6, I2 = 4.$$

Question 3

5



EXE 1 Question 8

(10 marks)

Since  $C(A) = 0$ , put A into S

$S = (A)$ ,  $\text{previous}(A) = \text{Nil}$

$S = (A)$

$C(B) = 0 + 11 = 11$ ,  $\text{previous}(B) = A$

Put B into S

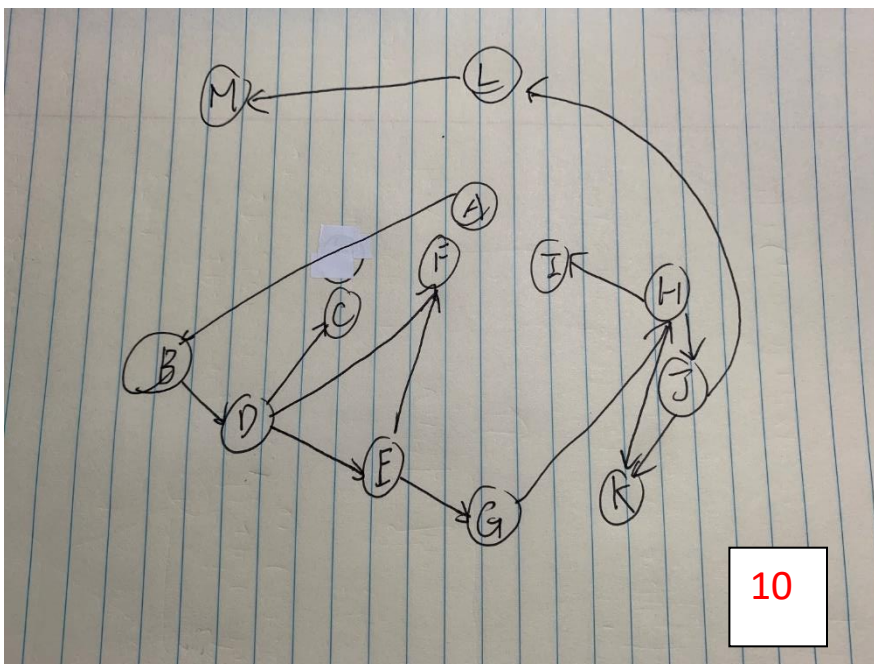
$S = (A, B)$

$S = (A, B)$

C and D are neighbors of B

$C(C) = 11 + 5 = 16$ ,  $\text{previous}(C) = B$

$C(D) = 11 + 23 = 34$ ,  $\text{previous}(D) = B$



C with the min cost

Put C into S

$S = (A, B, C)$

$S = (A, B, C)$

There are no neighbors to C

Then put D into S

$S = (A, B, C, D)$

$S = (A, B, C, D)$

E and F are neighbors to D

$C(E) = 34 + 5 = 39$ ,  $\text{previous}(E) = D$

$C(F) = 34 + 19 = 53$ ,  $\text{previous}(F) = D$

E with min cost

Put E into S

$S = (A, B, C, D, E)$

$S = (A, B, C, D, E)$

F and G are the neighbors of E

$C(F) = 39 + 9 = 48 < 53$ ,  $\text{previous}(F) = E$

$C(G) = 39 + 6 = 45$ ,  $\text{previous}(G) = E$

G with min cost

Then put G into S and update F

$S = (A, B, C, D, E, G)$

$S = (A, B, C, D, E, G)$

F and H are neighbors of G

$C(F) = 45 + 17 = 62$ ,  $\text{previous}(F) = G$

$C(H) = 45 + 27 = 72$ ,  $\text{previous}(H) = G$

Put F into S

Since F will repeat and become a loop

Put H into S

$S = (A, B, C, D, E, F, G, H)$

$S = (A, B, C, D, E, F, G, H)$

I and J are neighbors of H

$C(I) = 72 + 20 = 92$ ,  $\text{previous}(I) = H$

$C(J) = 72 + 2 = 74$ ,  $\text{previous}(J) = H$

J with min cost

Put J into S

$S = (A, B, C, D, E, F, G, H, J)$

$S = (A, B, C, D, E, F, G, H, J)$

K and L are neighbors of J

$C(K) = 74 + 12 = 86$ ,  $\text{previous}(K) = J$

$C(L) = 74 + 31 = 105$ ,  $\text{previous}(L) = J$

Put K into S

Since K don't have neighbors

Then put L into S

$S = (A, B, C, D, E, F, G, H, J, K, L)$

Therefore the shortest path from A to L is

$A \rightarrow B \rightarrow D \rightarrow E \rightarrow G \rightarrow H \rightarrow J \rightarrow L$

Question 4

EXE 1 Question 17

(10 marks)

Since  $C(1) = 0$ , put 1 into S

$S = (1)$ ,  $\text{previous}(1) = \text{Nil}$

$S = (1)$

$C(2) = 0 + 2 = 2$ ,  $\text{previous}(2) = 1$

$C(3) = 0 + 4 = 4$ ,  $\text{previous}(3) = 1$

2 is with min cost

Put 2 into S

$S = (1, 2)$

$S = (1, 2)$

3, 4 and 5 are neighbors to 2

$C(3) = 2 + 1 = 3 < 4$ ,  $\text{previous}(3) = 2$

$C(4) = 2 + 4 = 6$ ,  $\text{previous}(4) = 2$

$C(5) = 2 + 2 = 4$ ,  $\text{previous}(5) = 2$

3 is with min cost

Put 3 to S

$S = (1,2,3)$

$S = (1,2,3)$

5 is neighbors to 3

$C(5) = 3+3 = 6 > 4$ ,  $\text{previous}(5) = 3$

No update

5 is with min cost and put 5 into S

$S = (1,2,3,5)$

$S = (1,2,3,5)$

4 and 6 are neighbors of 5

$C(4) = 4+3 = 7 > 6$ ,  $\text{previous}(4) = 5$

$C(6) = 4+2 = 6$ ,  $\text{previous}(6) = 5$

4 with min cost and put 4 into S

$S = (1,2,3,4,5)$

$S = (1,2,3,4,5)$

6 is neighbors of 4

$C(6) = 6+2 = 8 > 6$ ,  $\text{previous}(6)=4$

6 is with min cost and put 6 into S

$S = (1,2,3,4,5,6)$

$S = (1,2,3,4,5,6)$

$C(2) = 2$ ,  $\text{previous}(2) = 1$

$C(3) = 3$ ,  $\text{previous}(3) = 2$

$C(4) = 6$ ,  $\text{previous}(4) = 2$

$C(5) = 4$ ,  $\text{previous}(5) = 2$

$C(6) = 6$ ,  $\text{previous}(6) = 5$

Therefore the shortest path from node 1 to 6 is

1 -> 2 -> 5 -> 6.

10