BDA SUMS - Bda sums

Big Data Analytics (University of Mumbai)

FEBRUARY 2020
M T W T F S S
1 2 3 4 5 6 7
8 9 10 11 12 13 14
15 16 17 18 19 20 21
22 23 24 25 26 27 28

Q. Apply Matrix - Matrix Multiplication using MapReduce Model and solve the following example.

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

→ consider A (i) 0 1 (j): $\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ B = 0 1 (j): $\begin{bmatrix} 3 & 4 \\ 3 & 4 \end{bmatrix}$ (k)

Step 1 :

matrix	i/k	j	value
A	0	0	1
A	0	1	2
A	1	0	1
A	1	1	2
B	0	0	3
B	0	1	4
B	1	0	3
B	1	1	4

Believe that life is worth living and your belief will help create the fact.

Step 2 : Map function (matrix, j, value)

i/k	value
A 0	(A, 0, 1)
A 0	(A, 1, 2)
A 1	(A, 0, 1)
A 1	(A, 1, 2)
B 0	(B, 0, 3)
B 0	(B, 1, 4)
B 1	(B, 0, 3)
B 1	(B, 1, 4)

Step 3 : Shuffle

i/k	value
A 0	(A, 0, 1)
A 0	(A, 1, 2)
A 1	(A, 0, 1)
A 1	(A, 1, 2)
B 0	(B, 0, 3)
B 0	(B, 0, 3)
B 1	(B, 1, 4)
B 1	(B, 1, 4)

Many of life's failures are people who didn't realize how close they were to success when they gave up.

Step : Shuffle / grouping

A B

(0, 0) ⟶ (A, 0, 1) (A, 1, 2)
(B, 0, 3) (B, 0, 3)

(0, 1) ⟶ (A, 0, 1) (A, 1, 2)
(B, 1, 4) (B, 1, 4)

(1, 0) ⟶ (A, 0, 1) (A, 1, 2)
(B, 0, 3) (B, 0, 3)

(1, 1) ⟶ (A, 0, 1) (A, 1, 2)
(B, 1, 4) (B, 1, 4)

Step 4 : Reduce function :

0 0 (1x3 + 2x3) = 9

0 1 (1x4 + 2x4) = 12

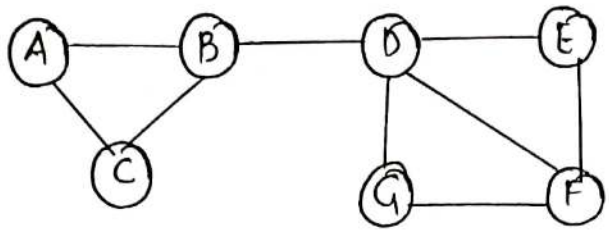
1 0 (1x3 + 2x3) = 9

1 1 (1x4 + 2x4) = 12

$\begin{bmatrix} 9 & 12 \\ 9 & 12 \end{bmatrix}$ is final ans. of matrix.

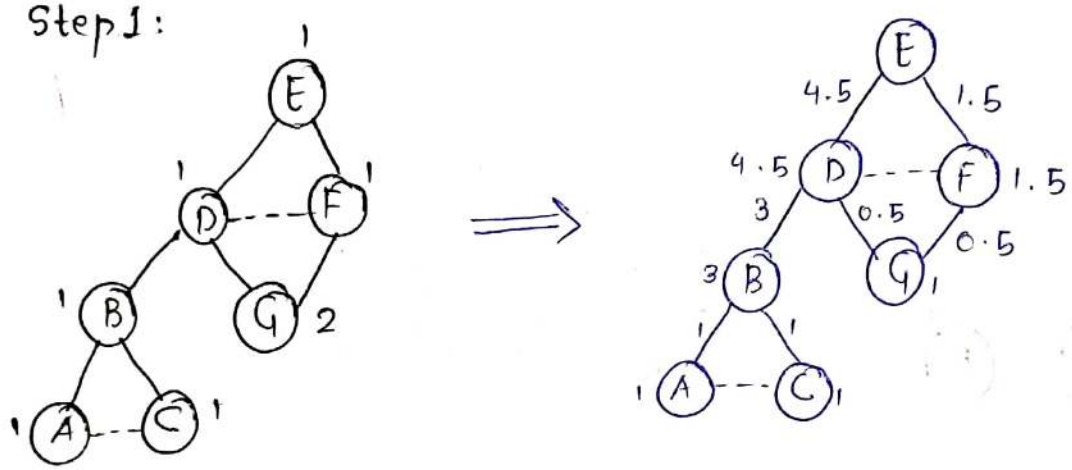
Successful people do what unsuccessful people are not willing to do. Don't wish it were easier; wish you were better.

Q. Figure is an Example of a social-network graph. Use the Girvan-Newman approach to find the betweenness of every edge.

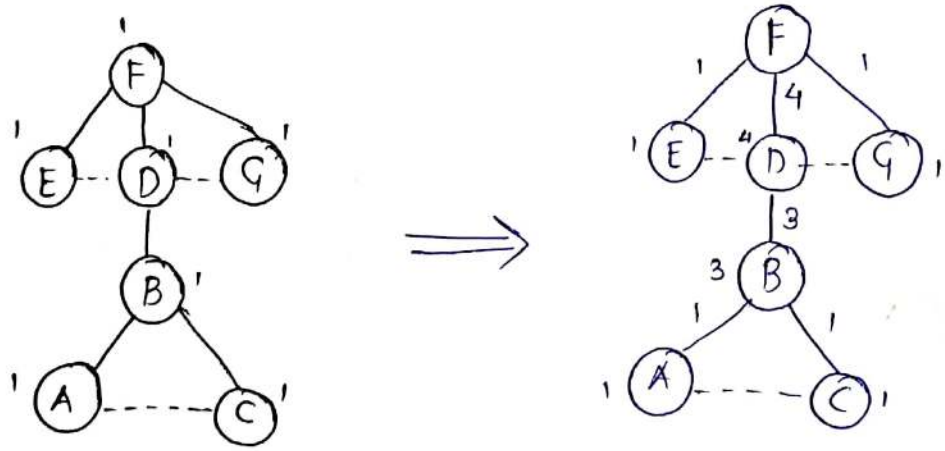


→ Solution:

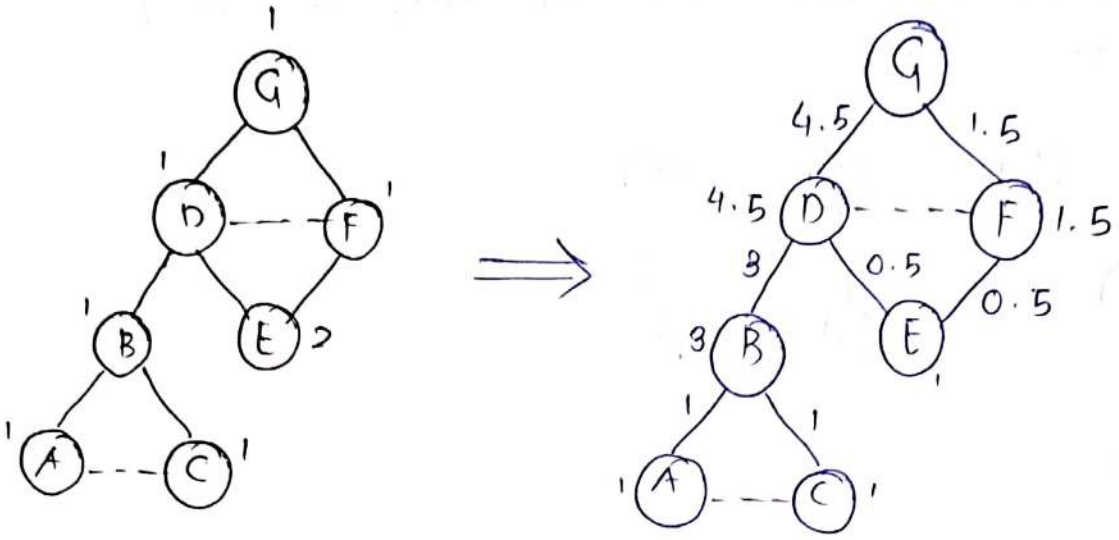
Step 1:



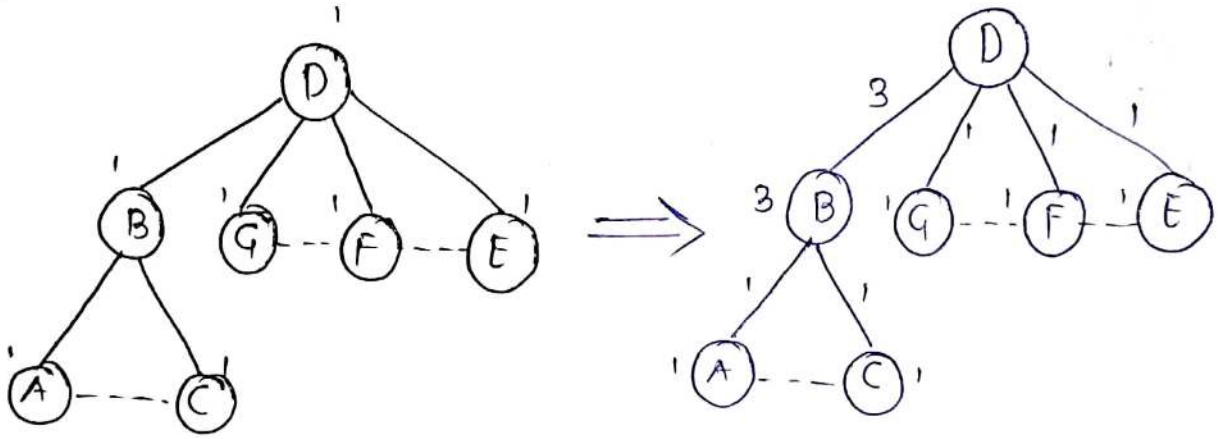
Step 2:



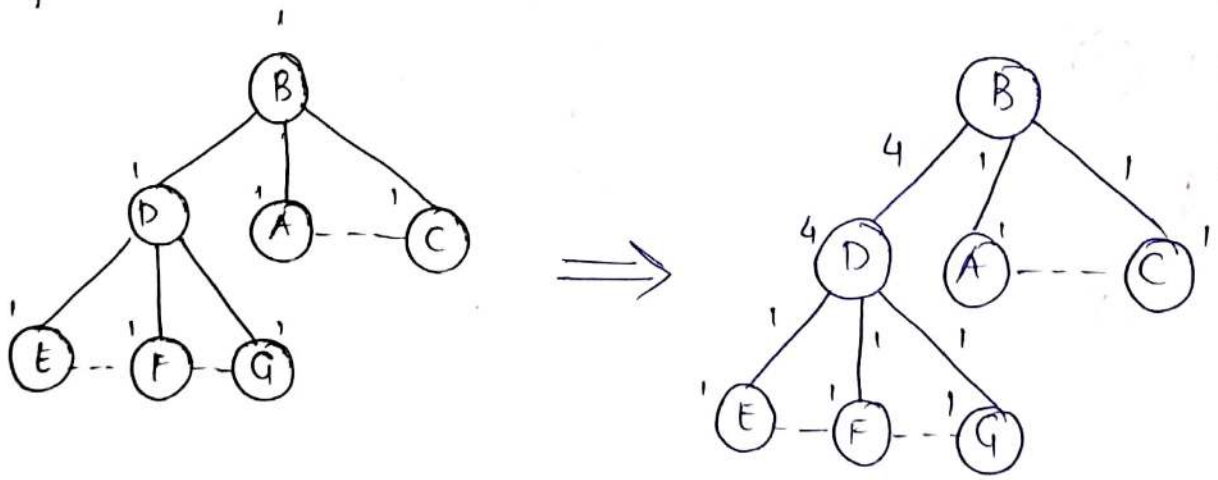
step 3 :



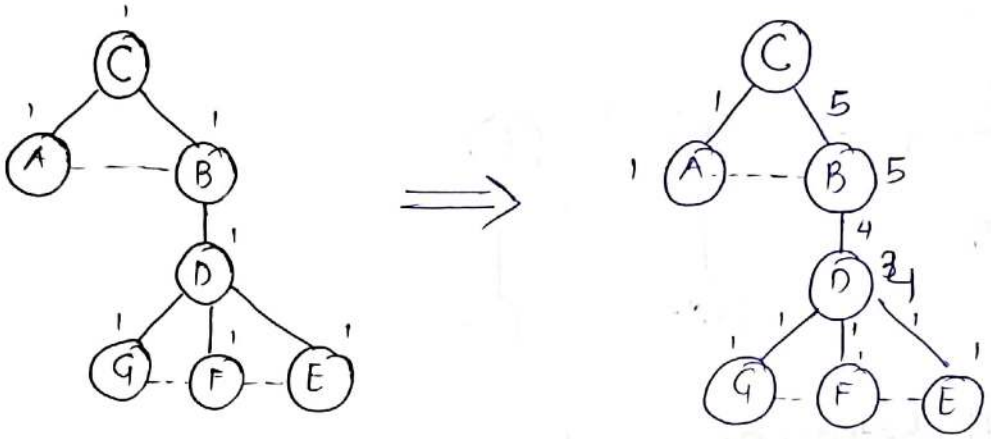
step 4 :



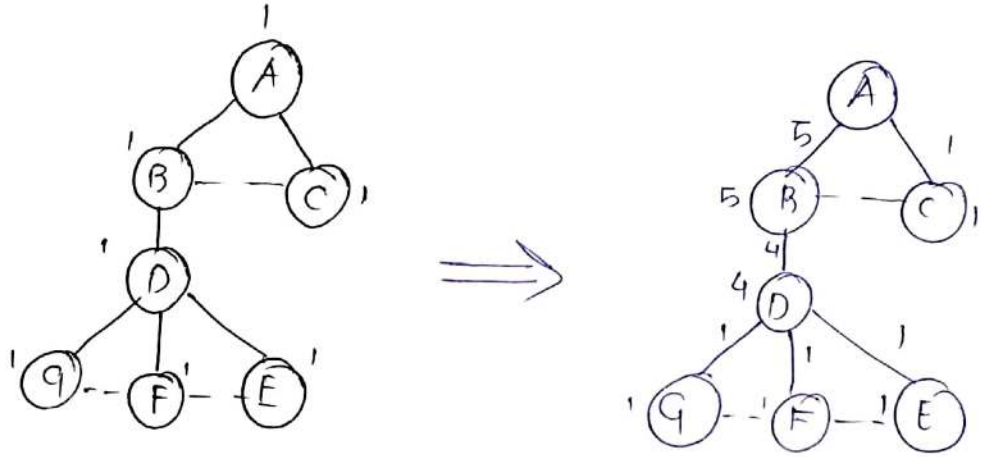
step 5 :



Step 6 :

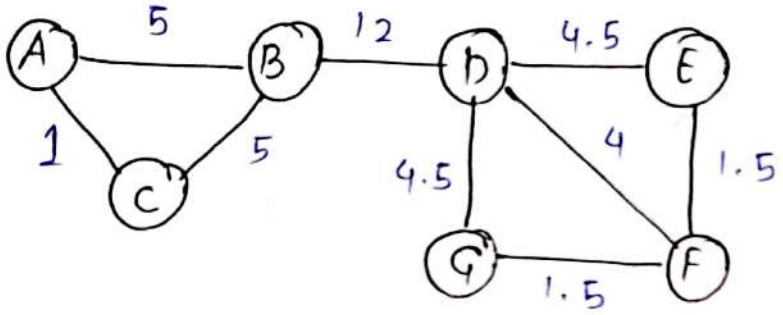


Step 7 :



Step 8 : Edges betweenness Calculation of every edge.

Edges	
AB	0+0+0+0+0+0+5 = 5
AC	0+0+0+0+0+0+1 = 1
BC	1+1+1+1+0+0+0 = 5
BD	0+0+0+0+4+4+4 = 12
DG	0.5+0+0+1+1+1+1 = 4.5
DE	0+0+0.5+1+1+1+1 = 4.5
DF	0+0+0+1+1+1+1 = 4
GF	0+0+1.5+0+0+0+0 = 1.5
EF	1.5+0+0+0+0+0+0 = 1.5



Graph coith betweeness value

4. Suppose a datastream consists of the integers:2 1 6 1 5 9 2 3 5,Let the function being used is

a)h(x)= 2x+3 mod 16

b)h(x)=4x+1 mod 16

c)h(x)=5x mod 16

Count Distinct Elements in a stream using FM algorithm.

Q. 2, 1, 6, 1, 5, 9, 2, 3, 5 ; Calculate distinct no. of elements in given data stream using FM algorithm

i) $h(x) = 2x+3 \bmod 16$
ii) $h(x) = 4x+1 \bmod 16$
iii) $5x \bmod 16$

i) $h(2) = 2(2)+3 \bmod 16 = 7$
$h(1) = 5$
$h(6) = 15$
$h(1) = 5$
$h(5) = 13$
$h(9) = 5$
$h(2) = 7$
$h(3) = 9$
$h(5) = 13$

Step 2:

$h(2) = 0111$
$h(1) = 0101$
$h(6) = 1111$
$h(1) = 0101$
$h(5) = 1101$
$h(9) = 0101$
$h(2) = 0111$
$h(3) = 1001$
$h(5) = 1101$

Step 3:

$h(2) = 0$
$h(1) = 0$
$h(6) = 0$
$h(1) = 0$
$h(5) = 0$
$h(9) = 0$
$h(2) = 0$
$h(3) = 0$
$h(5) = 0$

Step 4:

$r(a) = 0$

$\Rightarrow R = 2^r$

$= 2^0 = 1$

∴ There are 1 distinct elements for $h(x) = 2x+3 \bmod 16$

ii) $h(x) = 4x+1 \bmod 16$

$h(2) = 4(2)+1 \bmod 16 = 9$
$h(1) = 5$
$h(6) = 9$
$h(1) = 5$
$h(5) = 5$

$h(9) = 5$
$h(2) = 9$
$h(3) = 13$
$h(5) = 5$

Step 2 :

$h(2) = 1001$
$h(1) = 0101$
$h(6) = 1001$
$h(1) = 0101$
$h(5) = 0101$

$h(9) = 0101$
$h(2) = 1001$
$h(3) = 1101$
$h(5) = 0101$

Step 3 :

$h(2) = 0$
$h(1) = 0$
$h(6) = 0$
$h(1) = 0$
$h(5) = 0$

$h(9) = 0$
$h(2) = 0$
$h(3) = 0$
$h(5) = 0$

Step 4 :

$r(a) = 0$

$\Rightarrow R = 2^r \Rightarrow 2^0 = 1$

$\therefore$ There are/is 1 distinct elements for $h(x) = 4x+1 \bmod 16$.

iii) $5x \bmod 16$

$h(2) = 5(2) \bmod 16 = 10$
$h(1) = 5$
$h(6) = 14$
$h(1) = 5$
$h(5) = 9$

$h(9) = 13$
$h(2) = 10$
$h(3) = 15$
$h(5) = 9$

Step 2:

$h(2) = 1010$
$h(1) = 0101$
$h(6) = 1110$
$h(1) = 0101$
$h(5) = 1001$

$h(9) = 1101$
$h(2) = 1010$
$h(3) = 1111$
$h(5) = 1001$

Step 3:

$h(2) = 1$
$h(1) = 0$
$h(6) = 1$
$h(1) = 0$
$h(5) = 0$

$h(9) = 0$
$h(2) = 1$
$h(3) = 0$
$h(5) = 0$

Step 4:

$r(a) = 1$

$\Rightarrow R = 2^r \Rightarrow 2^1 = 2$

$\therefore$ There are 2 distinct elements for $h(x) = 5x \bmod 16$