DATA ANALYTICS

Assignment-4

P-Swetha
CSE-E
2107701277

Fm algorithm

i) $h(n) = (n * 17) \bmod 31$

ii) $h(n) = (n * 29) \bmod 37$

① 4,7, 4, 2, 8, 9, 7, 6, 2, 3, 8, 9, 10, 10

Element	Hash value	Binary Equivalent	Tail length(R)	No of distinct element (2^R)
4	6	00110	1	2
7	26	11010	1	2
2	3	00011	0	1
4	6	00110	1	2
8	12	01100	2	4
9	29	11101	0	1
7	26	11010	1	2
6	9	01001	0	1
2	3	00011	0	1
3	20	10100	2	4
8	12	01100	2	4
9	29	11101	0	1
10	15	01111	0	1
10	15	01111	0	1

or

(i) h(n) = (n * 17) mod 31

h(4) = 4 × 17 mod 31
= 6

h(7) = 7 × 17 mod 31 = 26

h(2) = 2 * 17 mod 31 = 3

h(2) = 3

h(8) = 8 * 17 mod 31 h(8) = 12

h(9) = 9 * 17 mod 31 h(9) = 29

h(6) = 6 * 17 mod 31

h(6) = 9

h(10) = 10 * 17 mod 31

h(10) = 15

5, 9, 15, 23, 42, 5

hash function

	hash value
	1
	0
1	1
5	
3	
5	
9	
15	

$$a = \frac{2+2+7+2+4+1+2+1+1+4+4+1+1+2}{16}$$

⇒ 27/14 = a = 1.928

(ii) h(n) = (n * 29) mod 37

h(4) = 4 × 29 mod 37

h(4) = 4 × 29 mod 37

h(4) = 5

h(7) = 7 * 29 mod 37

h(7) = 26

h(2) = 2 * 17 mod 31

h(2) = 3

h(4) = 6

h(8) = 12 → h(8) = 8 * 17 mod 31

h(9) = 9 * 17 mod 31 → h(8) = 12

h(9) = 9 * 17 mod 31

h(9) = 29

h(7) = 26

h(6) = 6 * 17 mod 31

h(6) = 9

h(2) = 3

h(3) = 3 * 17 mod 31

h(3) = 20

h(8) = 12

h(9) = 29

h(10) = 10 * 17 mod 31

→ h(10) = 15

5, 9, 15, 23, 42, 5, 9, 15, 30, 55

hash function $h(x) = (6x+1) \bmod 5$

Element	hash value	Binary Equivalent	Tail length (k)	No of distinct element (2^k)
5	1	00001	0	1
9	0	00000	1	2
15	1	00001	0	1
23	4	00100	2	4
42	3	00011	0	1
5	1	00001	0	1
9	0	00000	1	2
15	1	00001	0	1
30	1	00001	0	1
55	1	00001	0	1

$h(5) = 6(5)+1 \bmod 5$

$h(5) = 1$

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

$h(9) = 0$

$h(15) = 6(15)+1 \bmod 5$

$h(15) = 1$

$h(23) = 6(23)+1 \bmod 5$

$h(23) = 4$

$h(42) = 6(42)+1 \bmod 5$

$\rightarrow \boxed{h(42) = 3}$ distinct first element

$h(5) = 6(5)+1 \bmod 5$

$\rightarrow \boxed{h(5) = 1}$

$\boxed{h(9) = 0}$

$h(15) = 1$

$h(30) = 6(30)+1 \bmod 5$

$h(30) = 1$

$h(55) = 6(55)+1 \bmod 5 \rightarrow 1$

→ a ⇒ 2+2+1+2+4+1+2+1+1+4+1+1+1+6

14

⇒ 27/14

a = 1.928

h(6) = 6*2

h(6) = 26

h(2) = 21

h(3) = ?

(2) Given $h(x) = (x * 29) \bmod 37$:

Element	Hash value	Binary Equivalent	Tail length(R)	No of direct element(s)
4	5	00101	0	1
7	18	10010	1	2
2	21	10101	0	1
4	5	00101	0	1
10	10	01010	1	2
8	2	00010	1	2
9	18	10010	1	2
7	26	11010	0	1
6	21	10101	0	1
2	13	01101	1	2
3	10	01010	1	2
8	2	00010	0	1
9	31	11111	0	1
10	31	11111		1

$h(4) = 4 * 29 \bmod 37$

h(4) = 5

h(7) = 18

h(2) = 21

h(4) = 5

h(8) = 10

h(9) = 2

7) = 18

6) = 6 * 29 mod 37

h(6) = 26

(2) = 21

3) = 3 * 29 mod 37

h(3) = 13

h(8) = 10

h(9) = 2

h(10) = 10 * 29 mod 37

h(10) = 31

h(10) = 31

$$= \frac{1+2+1+2+2+2+1+1+2+2+1+1}{14}$$

b = 21/14

b = 1.5

	Avg of 2^R	median $\left(\frac{n+1}{2}\right)$
a	1.928	1.928
b	1.5	

distinct elements

$\Rightarrow 2^R$

ϕ

$\rightarrow \frac{1.928}{0.773} \Rightarrow \simeq 2$

$\frac{2+1}{2} \Rightarrow 3/2$

$\rightarrow 1.5$

$\simeq 1.$

(3) Insert element 10 and 7 in the bloom filter of size 5.

Consider these two hash functions

$h_1(x) = x \bmod 5$

$h_2(x) = (2x + 6) \bmod 5$

check the presence of elements 14 and 15

For 10: $h_1 = 0$ $h_2 = 1$

For 07: $h_1 = 2$ $h_2 = 0$

1	1	1	0	0

For 14: $h_1 = 4$ $h_2 = 4$ → At Index 4, the value is 0 which means that the element is not present in the set.

For 15, $h_1 = 0$ $h_2 = 1$ → the value is 1 at both 0 and 1 which mean that the element is not present.

Q4) You have bloom filter with the following specification,

→ (B) 1 Array size 1 10 bits

→ No of hash function 1?

h(4) = 5 h(4) = 5

→ Intially All bits are set to 0

$h1(x) = x \bmod 10$

$h(2)x) = (x \bmod 10+3) \bmod 10$

$h3(x) = (x \bmod 10+5) \bmod 10.$

Insert the element 1, 2, 2, 3 Also check the preference of element.

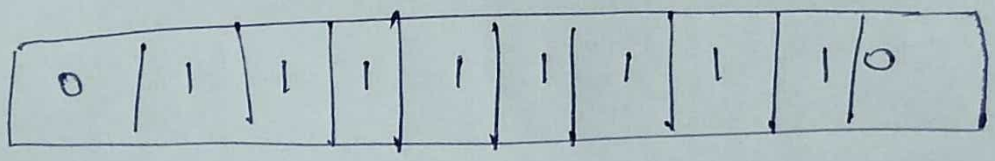
For 1: $h1=1$ $h2=4$ $h3=6$

For 2: $h1=1$ $h2=5$ $h3=7$

For 3: $h1=3$ $h2=6$ $h3=8$

After Inserting 1, 2, 2, 3

bloom filter looks like.



For 4:

$h1=4$ $h2=7$ $h3=9$

For Index: 4 : Value 1

Index 7 : Value 1

Index 9 : Value 0

Since the bit at index 4 is 0 the bloom filter indicates that the element is not present