

problem

PCY

$\min\_sup = 3$  ; Hash function =  $(i \times j) \bmod 10$

$$T_1 = \{1, 2, 3\} ; T_2 = \{2, 3, 4\} ; T_3 = \{3, 4, 5\}$$

$$T_4 = \{4, 5, 6\} ; T_5 = \{1, 3, 5\} ; T_6 = \{2, 4, 6\}$$

$$T_7 = \{1, 3, 4\} ; T_8 = \{2, 4, 5\} ; T_9 = \{3, 4, 6\}$$

$$T_{10} = \{1, 2, 4\} ; T_{11} = \{2, 3, 5\} ; T_{12} = \{3, 4, 6\}$$

$$\{1, 2, 3, 4, 5, 6\}$$

Sol : Index Count

1	→	4
2	→	6
3	→	8
4	→	9
5	→	5
6	→	4

$$C(2) = \frac{5+6}{2} = 5.5$$

$4! \times 2! \times 1!$

(combination)

$$T_1 = \{ \overset{1,2,3}{(1,2)}, \underline{(1,3)}, \underline{(2,3)} \} = (2, 3, 3)$$
$$T_2 = \overset{2,3,4}{\{ (2,3), (2,4), (3,4) \}} = (3, 4, 5)$$
$$T_3 = \overset{2,4,5}{\{ \cancel{(2,4)}, (3,5), (4,5) \}} = (\underline{5}, 3, 3)$$
$$T_4 = \{ \cancel{(4,5)}, \cancel{(4,6)}, (5,6) \} = (3, 4, \underline{1})$$
$$T_5 = \{ \underline{(1,3)}, (1,5), (3,5) \} = (3, 1, 3)$$
$$T_6 = \{ (2,4), (2,6), \underline{(4,6)} \} = (4, 1, 4)$$
$$T_7 = \{ \underline{(1,3)}, (1,4), (3,4) \} = (3, 2, \underline{5})$$
$$T_8 = \{ (2,4), (2,5), (4,5) \} = \underline{(4, 2, 3)}$$

$$T_9 = \{(3,4), (3,6), (4,6)\} = (5, 2, 4)$$

$$T_{10} = \{(1,2), (1,4), (2,4)\} = (2, 2, 4)$$

$$T_{11} = \{(2,3), (4,5), (3,5)\} = (3, 2, 3)$$

$$T_{12} = \{(3,4), (3,6), (4,6)\} = (5, 2, 4)$$

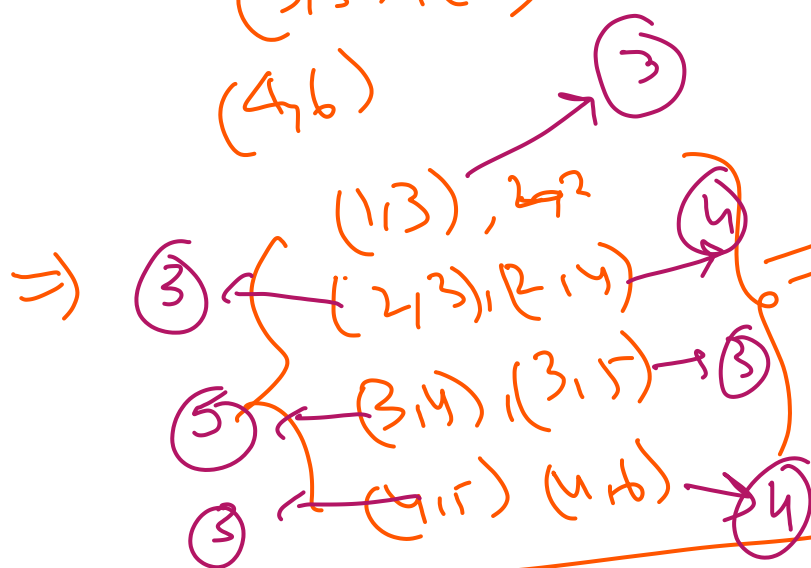
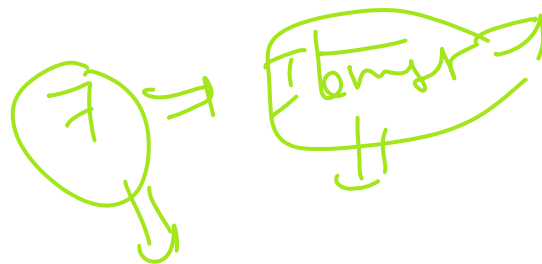
min\_sup  $\geq 3$ ,

$$\{(1,3), (2,3),$$

$$(2,3), (2,4), (3,4)$$

$$(3,5), (4,5)$$

$$(4,6)$$



$\geq$  min-support

Itemset  
frequent

Assume  $\Rightarrow$  Hash function is not provided.

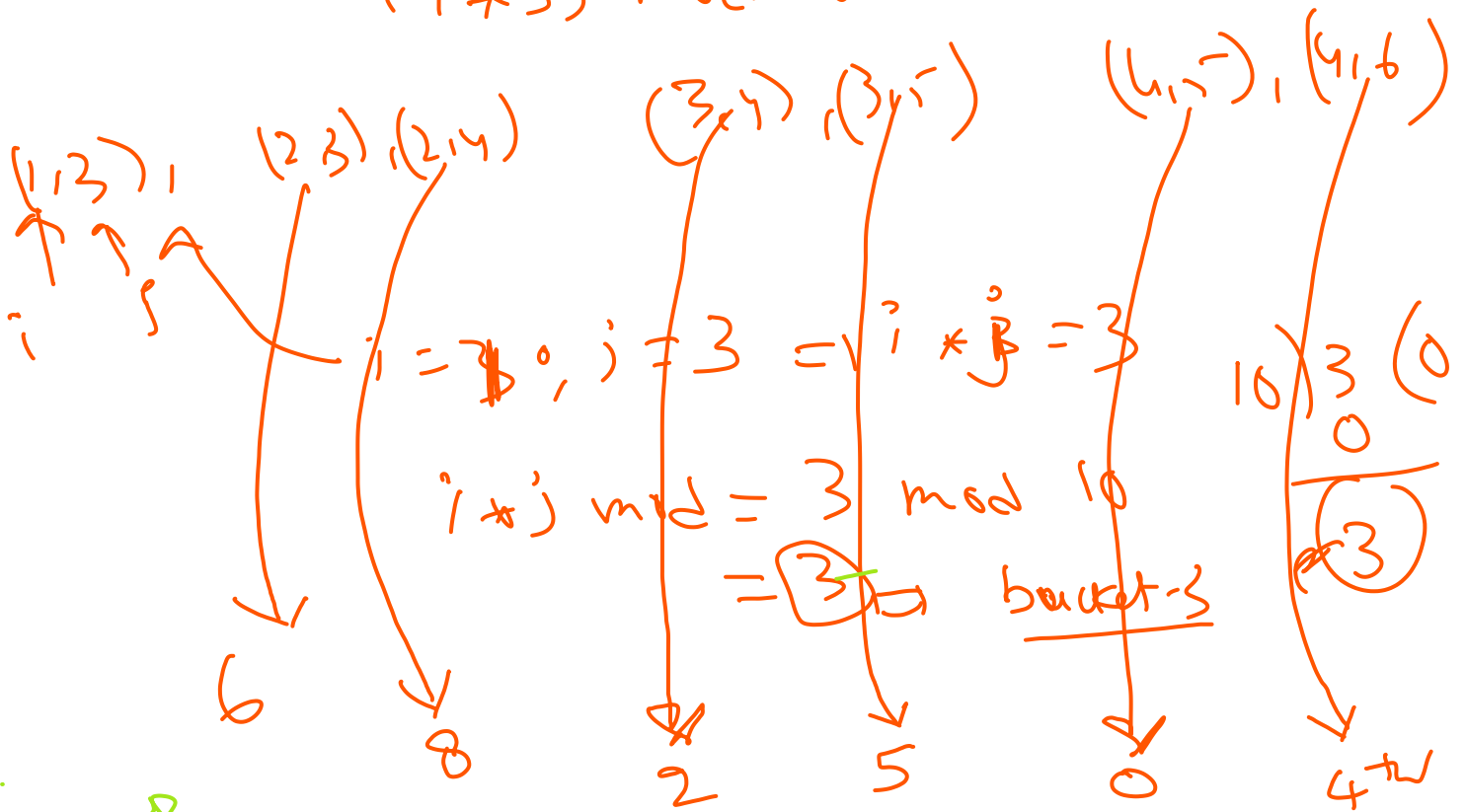
bucket

1	2	3	4	5	6	7	8	9	10	11	12
0	0	1, 1, 1	1, 1	1	0, 1	0	0	0	0	0	0

Itemset  $\Rightarrow$

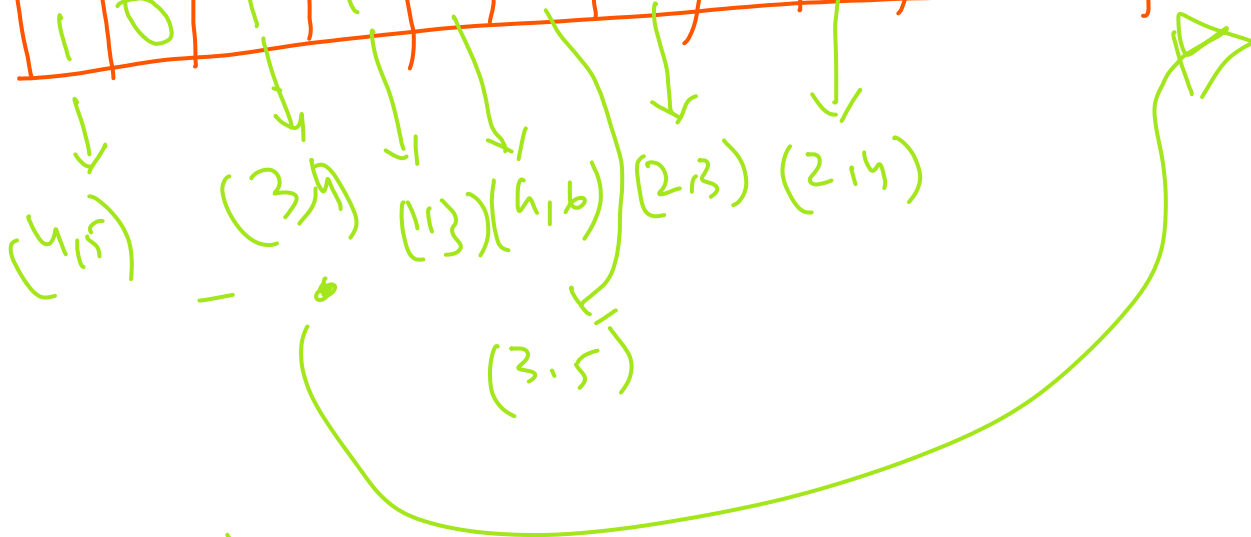
Assume: Hash function is given:

$$(i * j) \bmod 10$$



Bit map

0	1	2	3	4	5	6	7	8	9	10	11	12
1	0	1	1	1	1	1	0	1	0	0	0	0



over

# Multi stage PCY Algorithm

$$T_1 = \{A, B, D\}, T_2 = \{A, C, E\}, T_3 = \{A, D\}$$

$$T_4 = \{B, E\}, T_5 = \{A, C, D\}, T_6 = \{A, B, D, E\}$$

$$T_7 = \{B, D, E\}, T_8 = \{B, D\}$$

So:

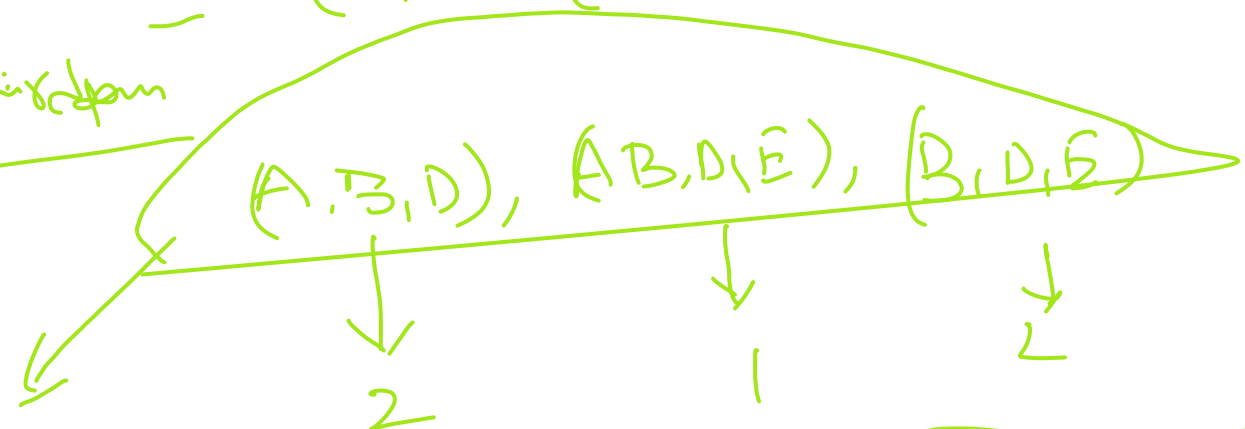
$$\begin{array}{l} \text{Item} = \{A, B, C, D, E\} \checkmark \Rightarrow 8 \times 0.5 = 4 \\ \downarrow \downarrow \downarrow \downarrow \downarrow \\ \text{count} = \{5, 5, 2, 6, 4\} \\ \checkmark \checkmark \times \checkmark \checkmark \Rightarrow \underline{\{A, B, D, E\}} \end{array}$$

## Doubletons

$$\begin{array}{cccccc} (A, B), & (A, D), & (A, E), & (B, D), & (B, E), & (D, E) \\ \downarrow \times & \downarrow & \downarrow \times & \downarrow & \downarrow & \downarrow \times \\ 2 & 4 & 2 & 4 & 3 & 2 \end{array}$$

$$(A, D), (B, D), (B, E)$$

## Threshold



$$\Rightarrow \boxed{C_2} \Rightarrow \boxed{\text{false positive}}$$

2 = mapReduce phase

~~1~~  $\rightarrow$  MapReduce ~~ch~~  
chunk  $\rightarrow$   $(AB, 1)$

Reduce  $\Rightarrow$   $(AB)$   
 $(BC)$  :

2nd phase  $\rightarrow$  MapReduce

