

Data Mining (CSE542)

Homework 04

ID: ____ Name: __조원석__ Date: __2023-04-04__

Task-1

Given the database in Table 8.2.

- (a) Using $minsup = 3/8$, show how the Apriori algorithm enumerates all frequent patterns from this dataset.
- (b) With $minsup = 2/8$, show how FPGrowth enumerates the frequent itemsets.

Table 8.2. Transaction database

tid	itemset
t_1	<i>ABCD</i>
t_2	<i>ACDF</i>
t_3	<i>ACDEG</i>
t_4	<i>ABDF</i>
t_5	<i>BCG</i>
t_6	<i>DFG</i>
t_7	<i>ABG</i>
t_8	<i>CDFG</i>

(a) Using $\text{minsup} = 3/8$, show how the Apriori algorithm enumerates all frequent patterns from this dataset.

Level one's Frequent pattern

A : 5, B : 4, C : 5, D : 6, F : 4, G : 5

⇒ A, B, C, D, F, G

Level two's Frequent pattern

AB : 3, AC : 3, AD : 4, AF : 2, AG : 2, BC : 2, BD : 2, BG : 2, CD : 4, CF : 3, CG : 3, DF : 4, DG : 3, FG : 3

⇒ AB, AC, AD, CD, CF, CG, DF, DG, FG

Level three's Frequent pattern

ABC : 1, ABD : 2, **ACD : 3, CDF : 3, CDG : 2, CFG : 2, DFG : 3**

⇒ ACD, CDF, DFG

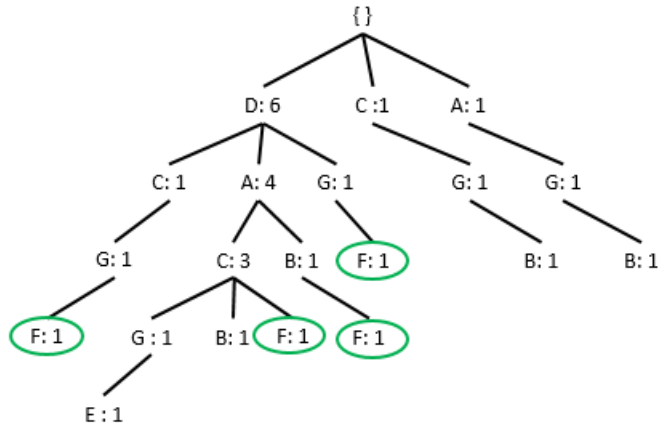
Level four's and five's Frequent pattern

ABCD : 1, ABDF : 1, ACDE : 1, ACDF : 1, ACDG : 1, ADEG : 1, CDEG : 1, CDFG : 1, ACDEG : 1

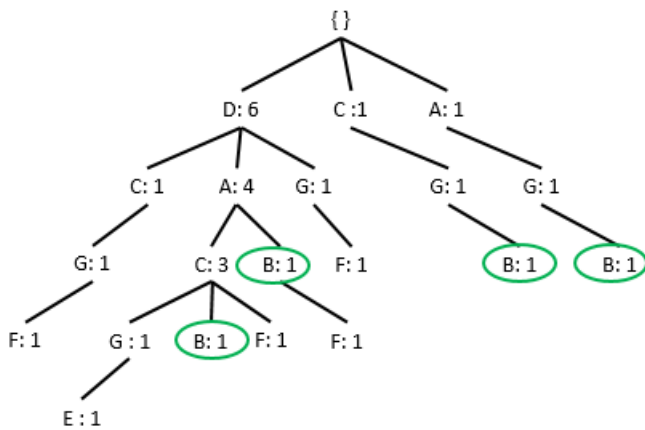
(b) With minsup = 2/8, show how FPGrowth enumerates the frequent itemsets.k

{D:6}, {A:5}, {C:5}, {G:5}, {B:4}, {F:4}

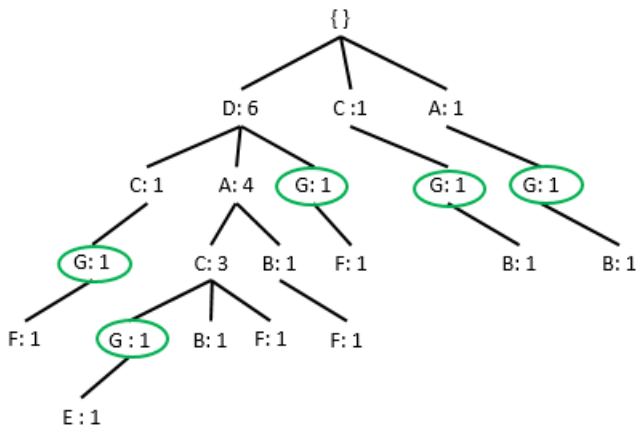
F : {DF: 4}, {AF: 2}, {GF: 2}, {CF: 2}, {DAF: 2}, {DGF: 2}, {DCF: 2}



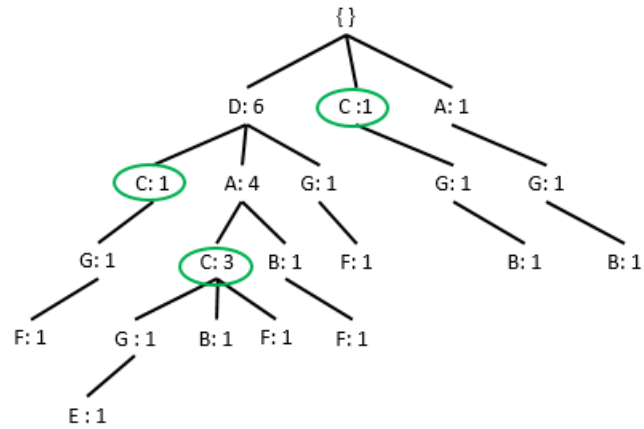
B : {AB: 3}, {BC: 2}, {BD: 2}, {BG: 2}, {ABD: 2}



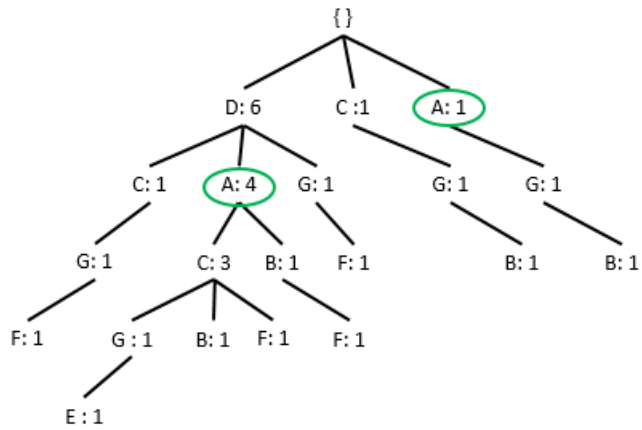
G : {CG: 3}, {DCG: 2}, {AG: 2}, {DG: 3}



C : {AC: 3}, {DAC: 3}, {DC: 4}



A: {DA: 4}



T1 = DACB => X

T2 = DACF => X

T3 = DACGE => X

T4 = DABF => X

T5 = CGB => X

T6 = AGB => X

T7 = AGB => X

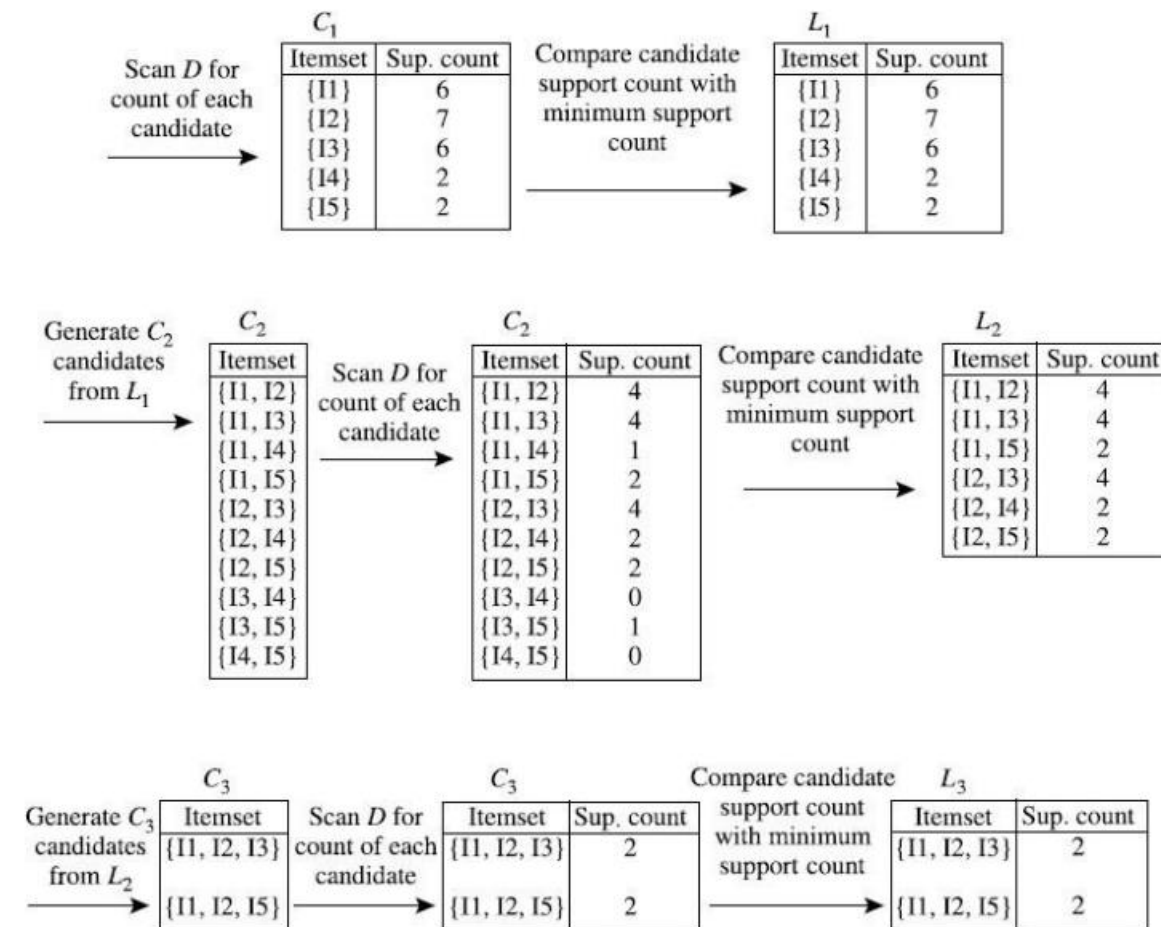
T8 = DCGF => X

An example is solved for Apriori and FPGrowth Algorithms

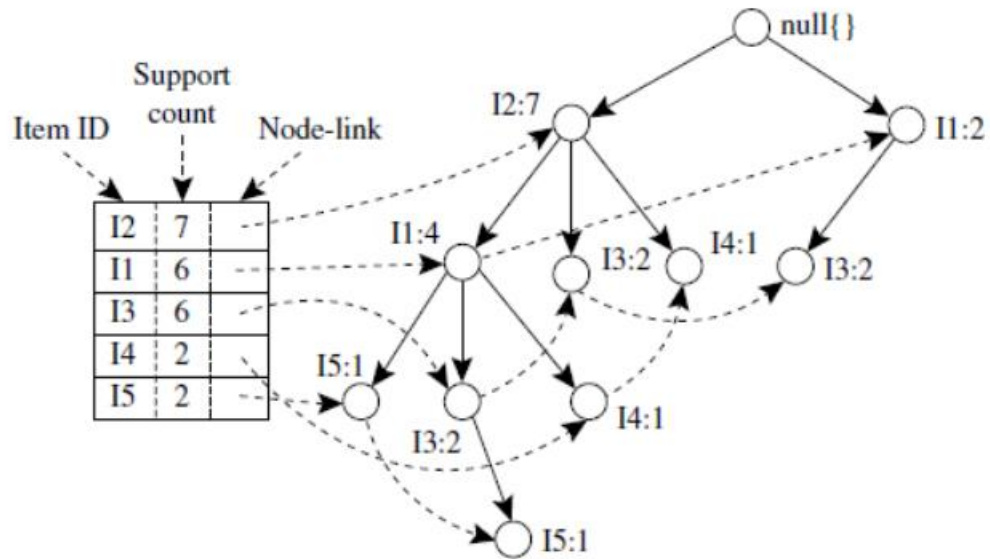
Dataset D

<i>TID</i>	<i>List of item_IDs</i>
T100	I1, I2, I5
T200	I2, I4
T300	I2, I3
T400	I1, I2, I4
T500	I1, I3
T600	I2, I3
T700	I1, I3
T800	I1, I2, I3, I5
T900	I1, I2, I3

Frequent Itemsets through Apriori Algorithm



Frequent Itemsets through FPgrowth Algorithm



Item	Conditional Pattern Base	Conditional FP-tree	Frequent Patterns Generated
I5	{{I2, I1: 1}, {I2, I1, I3: 1}}	$\langle I2: 2, I1: 2 \rangle$	{I2, I5: 2}, {I1, I5: 2}, {I2, I1, I5: 2}
I4	{{I2, I1: 1}, {I2: 1}}	$\langle I2: 2 \rangle$	{I2, I4: 2}
I3	{{I2, I1: 2}, {I2: 2}, {I1: 2}}	$\langle I2: 4, I1: 2 \rangle, \langle I1: 2 \rangle$	{I2, I3: 4}, {I1, I3: 4}, {I2, I1, I3: 2}
I1	{{I2: 4}}	$\langle I2: 4 \rangle$	{I2, I1: 4}