

004.93

[1, 2].

(),

« , ».

$$D(1): D = \{T_1, T_2, \dots, T_{N_D}\}, \tag{1}$$

[3], T_j ,

$j = 1, 2, \dots, N_T$

$N_D = |D|$ ó

() D .

T_j (2):

$$T_j = (tid_j, item_j), \tag{2}$$

tid_j ó

j -

$$T_j; item_j = \{t_{1j}, t_{2j}, \dots, t_{N_{item_j}j}\} \subseteq I \text{ ó}$$

$T_j; t_{ij}$ ó i - $item_j, i = 1, 2, \dots, N_{item_j};$

$N_{item_j} = |item_j|$ ó

$$item_j; I = \{\tau_1, \tau_2, \dots, \tau_{N_I}\} \text{ ó}$$

$item_j$

$T_j, j = 1, 2, \dots, N_T$

$D; \tau_a$ ó -

$I,$

$a = 1, 2, \dots, N_I; N_I = |I|$ ó

$I.$

T_j

(3):

$$T_j = item_j = \{t_{1j}, t_{2j}, \dots, t_{N_{Ij}j}\}, \tag{3}$$

t_{ij} ó i -

j -

T_j

(4):

$T_j;$

(4)

$$t_{ij} = \begin{cases} 1, \\ 0, \end{cases}$$

T_j

D

$item_j,$

$I.$

AR

$X \rightarrow Y,$

$X \ Y$

(5) [365]:

$$X \rightarrow Y: X \subset I, Y \subset I, X \cap Y = \emptyset. \tag{5}$$

(5)

..

: "

X

Y" " X, Y" [3].

" \emptyset - : 9 , , "
 , 2012.
 AR
 $D ($)
 $\tau_a \in I, a = 1, 2, \dots, N_I$ [4].

(6) [365]: $X \subset I$ D $\text{supp}(X)$,

$$\text{supp}(X) = \frac{N_{T \in D | X \subseteq T}}{N_D}, \tag{6}$$

$N_{T \in D | X \subseteq T}$ \acute{o} T D ,
 X X D ,
 $\text{supp}(X)$
 minsupport, [3, 4].

(7): $\text{supp}(X \rightarrow Y)$ $X \rightarrow Y$ $X \cup Y$

$$\text{supp}(X \rightarrow Y) = \text{supp}(X \cup Y). \tag{7}$$

$X \rightarrow Y$
 [365].
 $\text{conf}(X \rightarrow Y)$ $X \rightarrow Y$ (8) [4, 5]:

$$\text{conf}(X \rightarrow Y) = \frac{\text{supp}(X \cup Y)}{\text{supp}(X)}. \tag{1.8}$$

()
 $\text{minconfidence}(X \rightarrow Y)$.

D , \acute{o} [365]. [365]:

\acute{o} X
 $\text{minsupport}(X)$,
 $X \subset I$;
 \acute{o} $X \rightarrow Y$
 $\text{minconfidence}(X \rightarrow Y)$.

$$D = \{T_1, T_2, \dots, T_{N_T}\},$$

 [3, 4]. . .

ϵ_I .

[367]:

$$X \rightarrow Y \quad \text{(Piatetsky-Shapiro).} \quad (9): \quad (9)$$

$$\text{supp}(X \rightarrow Y) \approx \text{supp}(X)\text{supp}(Y).$$

$$\frac{\text{supp}(X \rightarrow Y)}{\text{supp}(X)\text{supp}(Y)} > 1, \quad X \quad Y$$

$$\frac{\text{supp}(X \rightarrow Y)}{\text{supp}(X)\text{supp}(Y)} < 1, \quad Y \quad X$$

[6];

$$(10): \quad X \rightarrow Y$$

$$\left| \frac{\text{supp}(X \rightarrow Y)}{\text{supp}(X)\text{supp}(Y)} - 1 \right| \geq \varepsilon_I, \quad (10)$$

$$\frac{\text{supp}(X \rightarrow Y)}{\text{supp}(X)\text{supp}(Y)} - 1 \geq \varepsilon_I \quad X \rightarrow Y$$

$$-\left(\frac{\text{supp}(X \rightarrow Y)}{\text{supp}(X)\text{supp}(Y)} - 1 \right) \geq \varepsilon_I \quad X \rightarrow \bar{Y} \quad [7].$$

$$X \rightarrow Y, \quad (11):$$

$$\begin{cases} \text{supp}(X \rightarrow Y) \geq \text{minsupport}; \\ \text{conf}(X \rightarrow Y) \geq \text{minconfidence}; \\ \left| \text{supp}(X \rightarrow Y) / \text{supp}(X)\text{supp}(Y) - 1 \right| \geq \varepsilon_I. \end{cases} \quad (11)$$

$$D = \{T_1, T_2, \dots, T_{N_T}\}, \quad I = \{\tau_1, \tau_2, \dots, \tau_{N_I}\};$$

$$T_j \subseteq I, \quad \text{minsupport}(\quad),$$

$$\text{minconfidence}(X \rightarrow Y) \quad \text{minsupport}(\quad) \quad \text{minconfidence}(X \rightarrow Y)$$

$$\varepsilon_I \cdot \quad (\quad)$$

[3610].

Support-Confidence Framework) [3, 8] « - » (SCF,

[8]:

$$(\quad)$$

$$: \text{supp}(X) \geq \text{minsupport};$$

$$A,$$

" \emptyset - , 2012. : 9 , , " $X \rightarrow Y$: $\forall Y \subset A , X = A - Y ,$
conf($X \rightarrow Y$) \geq minconfidence .

(, ,) ,
minsupport minconfidence ,

()
(SETM, Set-oriented mining)
« SQL »
[3, 6, 9].

[9].
SETM
Apriori ()

[367, 10].

[10]:
() .
t- t-
t- t-
t- t-

$X \subseteq I \quad Y \subseteq I , \quad X \subseteq Y .$
[367]:
 $Y \subseteq I \quad X \subseteq Y$

(12):
 $X \subseteq Y : \text{supp}(Y) \leq \text{supp}(X) .$ (12)
 $Y \subseteq I$
 $X \subseteq Y .$

Apriori ,
[10].
X
 $\text{supp}(X) < \text{minsupport} .$ t-

" \emptyset - , 2012. : 9 , , "

[10].

Apriori $D(1)$,

[3, 6, 10]. , D

Apriori

$D(1)$

Apriori:

AprioriTID AprioriHybrid [367].

AprioriTID , $D(1)$

[367].

(,) , .

D ,

Apriori AprioriTID

[4, 6, 7].

),

D ,

AprioriTID

AprioriHybrid

Apriori AprioriTID , [365].

Apriori AprioriTID

t -

t -

AprioriHybrid

Apriori AprioriTID.

(DHP, Direct Hashing and Pruning)

(

) [3, 5, 11].

t -

t -

[11].

t -

t -

" \emptyset - , 2012. : 9 , , " 39

Apriori, DHP , - ,

minsupport [11].

, DHP .

H_2 , - .

C_t - H_t , t - ,

Apriori, DHP , AprioriTID [11].

AprioriTID

AprioriHybrid, DHP

DHP

Partition () $D(1)$

[12].

Partition Apriori,

[3, 12].

Partition

[12].

$D(1)$, Apriori [365, 10],

Eclat

1997 M. Zaki [6, 13].

Eclat

Apriori [13]. Eclat

Apriori, [13].

Eclat

Apriori [10, 13].

Partition.

Apriori Eclat

Hybrid, Apriori, Eclat.

Apriori

Eclat.

FPG (Frequent Pattern Growth,)

[567] Eclat

FPG $D(1)$. $D(1)$

ó
 (FPG).
 (t - [6, 7];
 ó (): $\text{supp}(X)$
 tid_j T_j ,
 $D[3, 4, 7];$
 $D(1)$ X ,
 τ_a , $\tau_a \in I$
 1 (k ó
).

Apriori	-			k
AprioriTid	-			1

Apriori Hybrid	-			k
DHP	-			1
Partition				k
Eclat				1
Hybrid		/	/	k
FPG				1

[3, 6, 7, 14]:

ó
 ó

ó
 ó

ó

0111U000059),

04922 «

ö (,
 ö (.

1. Engelbrecht A. Computational intelligence: an introduction / A. Engelbrecht. ó Sidney : John Wiley & Sons, 2007. ó 597 p.

2. Shin Y.C. Intelligent systems : modeling, optimization, and control / C. Y. Shin, C. Xu. ó Boca Raton: CRC Press, 2009. ó 456 p.
3. Zhang C. Association rule mining: models and algorithms / C. Zhang, S. Zhang. ó Berlin : Springer-Verlag. ó 2002. ó 238 p.
4. Gkoulalas-Divanis A. Association Rule Hiding for Data Mining / A. Gkoulalas-Divanis, V. S. Verykios. ó New York : Springer-Verlag. ó 2010. ó 150 p.
5. Zhao Y. Post-mining of association rules: techniques for effective knowledge extraction / Y. Zhao, C. Zhang, L. Cao. ó New York : Information Science Reference. ó 2009. ó 372 p.
6. J.-M. Adamo. Data mining for association rules and sequential patterns: sequential and parallel algorithms / Adamo J.-M. ó New York : Springer-Verlag. ó 2001. ó 259 p.
7. Koh Y. S. Rare Association Rule Mining and Knowledge Discovery / Y. S. Koh, N. Rountree. ó New York : Information Science Reference. ó 2009. ó 320 p.
8. Agrawal R. Mining association rules between sets of items in large databases / R. Agrawal, T. Imielinski, A. Swami // Management of Data : International Conference ACM SIGMOD, Washington, 26-28 May 1993 : Proceedings of the Conference. ó New York : ACM Press, 1993. ó P. 207-216.
9. Houtsma M. Set-oriented mining for association rules in relational databases / M. Houtsma, A. Swami // Data Engineering : The 11th International Conference, Taipei, Taiwan, 6-10 March 1995 : Proceedings of the Conference. ó Washington : IEEE Computer Society, 1995. ó P. 25-33.
10. Agrawal R. Fast algorithms for mining association rules / R. Agrawal, R. Srikant // Very Large Data Bases: The 20th International Conference VLDB'94, Santiago de Chile, Chile, 12-15 September 1994 : Proceedings of the Conference. ó San Francisco : Morgan Kaufmann, 1994. ó P. 487-499.
11. Park J. S. An effective hash based algorithm for mining association rules / J. S. Park, M. Chen, P. S. Yu // Management of Data : International Conference ACM SIGMOD, San Jose, California, 22-25 May 1995 : Proceedings of the Conference. ó New York : ACM Press, 1995. ó Volume 24, Issue 2. ó P. 175-186.
12. Savasere. A. An efficient algorithm for mining association rules in large databases / A. Savasere, E. Omiecinski, S. Navathe // Very Large Data Bases : The 21th International Conference VLDB '95, Zurich, Switzerland, 11-15 September 1995 : Proceedings of the Conference. ó San Francisco : Morgan Kaufmann, 1995. ó P. 432-444.
13. Zaki M. J. New Algorithms for Fast Discovery of Association Rules / M. J. Zaki, S. Parthasarathy, M. Ogihara, W. Li // Knowledge Discovery and Data Mining : The 3rd International Conference KDD-97, Newport Beach, California, 14-17 August 1997 : Proceedings of the Conference. ó Palo Alto, California : AAAI Press, 1997. ó P. 283-286.
14. . . . / . . . // X I : 16- , 17619 2012 . : . ó : , 2012. ó .6. ó .216 22.