TELKOMNIKA, Vol. 11, No. 2, February 2013, pp. 941~947
ISSN: 2302-4046

An Approach for Briefest Rules Extraction Based On Compact Dependencies

Yang Kai*[1,3], Jin Yong-long[2], He Zhi-jun[3], Ma Yuan[1]
[1]School of Software, University of Science and Technology LiaoNing, 114051, China
[2]Technical Center of TangShan Iron & Steel, HeBei Iron & Steel Group, HeBei, 063016,China
[3]School of materials and metallurgy, University of Science and Technology LiaoNing, 114051, China
Corresponding author, e-mail: asyangkai@126.com, yonglongjin@126.com, hzhj2002@126.com,mayuanas@sina.com

Abstract

Concept lattice is an effective tool for knowledge representation and data analysis. It have been successfully applied to many fields. This paper illustrates the theory of compact dependencies in concept lattice, then study an approach for briefest rules extraction based on compact dependencies, including briefest decision rules and briefest association rules with confidence of 1. Finally, the method is applied to the example and the relationship between many process parameters and aim parameters,as well as process parameters and other process parameters, are discussed. It provides a useful decision support tool for advanced production management and increases the ability of the decision, which accomplished the validity of the enterprises.

Keywords: *Concept lattice, Value Dependency, Decision rules, Association Rules*

1. Introduction

The theory of value dependency is the common concern of many fields, such as database theory, rough set theory, formal concept analysis and data mining [1]. Discovery of dependencies existing in an instance of a relation received considerable interest as it allowed automatic database analysis. Some areas, such as health care [2], intrusion detection [3], database management [4], graph theory [5], artificial intelligence [6] are among the main applications benefiting from efficient dependencies discovery algorithms and data mining algorightms.

With the development of a formal context, the number of value dependencies that hold in the context increases exponentially. How to automatically deal with the huge information has become a very important research topic. In 1986 Guigues and Duquenne put forth the theory of Guigues-Duquenne dependency basis that bears their name [7]. The dependency basis is a set of value dependenies that holds in a given formal context and may satisfied the two follow conditions: all the value dependencies hold in the context can be derived from the dependency basis according to armstrong axiomatic system, while each value dependency in the dependency basis can not be derived from other value dependencies in the set according armstrong axiomatic system. To find out the dependency basis and explore their application, A lot of research work has been carried out in recent years and many approachs have been applied to the reduction of association rules [8-11]. In [12], authors give a thorough formal study of the related inference mechanisms allowing to derive all redundant association rules starting from succinct ones. The concept of compact dependencies is firstly presented in [1]. It has been proved that all the value dependencies hold in a given context can be derived completely with no redundant by means of compact dependencies. Furthermore, the compact dependencies is obtained more convenient and quickly than GD basis. For their unique excellent concise character, compact dependencies will be widely used in many fields, such as briefest rules extraction, personalization services and so on.

To make the best of the great amount of production data accumulated in the course of industry system development, and to satisfy the needs of production and supervision management at each level, the construction of decision support system is of the greatest importance. To analyse the inherence relations between production parameters, it is proposed

Received September 15, 2012; Revised December 31, 2012; Accepted January 15, 2013

to extract some laws from lots of production data using the data mining technology. This paper introduce the theory of compact dependencies in concept lattice, then propose an approach for briefest rules extraction based on compact dependencies, including briefest decision rules and briefest association rules with confidence of 1. In the last, the paper describes their applications on a decision Table and comments the corresponding results.

2. Fundamental Definition and Theorem

Before proceeding, we briefly recall the FCA terminology [13].

Definition 1 Given a formal context $K=(U,M,I)$, where U is called *a set of objects*, M is called *a set of attributes*, and the binary relation $I\subseteq U\times M$ specifies which objects have which attributes, the derivation operators $f(\cdot)$ and $g(\cdot)$ are defined for $A\subseteq U$ and $B\subseteq M$ as follows:

$$f(A)=\{m\in M\mid \forall u\in A,(u,m)\in I\};\quad g(B)=\{u\in U\mid \forall m\in B,(u,m)\in I\}.$$

In words, $f(A)$ is the set of attributes common to all objects of A and $g(B)$ is the set of objects sharing all attributes of B. A formal concept of the context (U,M,I) is a pair (A,B), where $A\subseteq U, B\subseteq M$, $f(A)=B$ and $g(B)=A$. The set A is called the *extent* and B is called the *intent* of the concept (A,B).

Definition 2 A concept (A_1,B_1) is a *subconcept* of (A_2,B_2) if $A_1\subseteq A_2$ (equivalently, $B_2\subseteq B_1$). In this case, (A_2,B_2) is called a *superconcept* of (A_1,B_1). We write as $(A_1,B_1)\le(A_2,B_2)$. If $(A_1,B_1)\le(A_2,B_2)$ and there is no (C,D) such that $(A_1,B_1)\le(C,D)\le(A_2,B_2)$, then (A_1,B_1) is a *direct subconcept* of (A_2,B_2) and (A_2,B_2) is a *direct superconcept* of (A_1,B_1). The set of all concepts ordered by $\le$ forms a lattice, which is denoted by $\mathrm{B}(K)$ and called the *concept lattice* of the context *K*.

Definition 3 Let (U,M,I) be a formal context. If $X,Y\subseteq M$, formulas $X\to Y$ are called *value dependencies*, in which X, Y are called *front component* and *back component* respectively. If value dependencies $X\to Y$ satisfy $g(X)\subseteq g(Y)$, we call that $X\to Y$ *holds* in the formal context, i.e. in data Tables with binary attributes, and have the following meaning: Each object having all attributes from X has also all attributes from Y.

Definition 4 For arbitrary two concept (A,B) and (A_1,B_1) such that (A,B) is a direct subconcept of (A_1,B_1) (or (A_1,B_1) is a direct superconcept of (A,B)), we define attributes set $B/B_1=\{m\mid m\in B\wedge m\notin B_1\}$, which is called *intent waned-value* of concept (A,B). All the intent waned-value of concept (A,B) are denoted by $\vee(A,B)$. We set $P(A,B)=\bigcup\vee(A,B)$ in which $\bigcup\vee(A,B)$ means union of all elements in $\vee(A,B)$。

Example 1. For concept #13 in Figure1, $\vee(14,b1c1d1e2)=\{c1d1e2,b1d1e2,b1c1d1\}$, $\mathrm{P}(14,b1c1d1e2)=\{b1,c1,d1,e2\}$.

Definition 5 Let (A,B) be a formal concept, a hypergraph $\mathrm{H}(A,B)$ is defined as follows. The vertex set of $\mathrm{H}(A,B)$ is $\mathrm{P}(A,B)$ and the edge set is $\mathrm{H}(A,B)$. The hypergraph $\mathrm{H}(A,B)$ is called *waned-value hypergraph* of concept (A,B). Let $p\subseteq\mathrm{P}(A,B)$ such that $\forall v\in\vee(A,B), v\bigcap p\neq\phi$, then p is called a *transversal* of $\mathrm{H}(A,B)$. If p is a transversal of $\mathrm{H}(A,B)$ such that $p_1\subset p$ and p_1 is not a transversal of $\mathrm{H}(A,B)$, we call p a *minnimal transversal* of $\mathrm{H}(A,B)$. *Example 2.* The minimal transversals of concept $\mathrm{H}(14,b1c1d1e2)$ in Figure1 are $\{c1b1, c1e2, d1, e2b1\}$.

3. Compact Dependencies and Briefest Rules

Let $S=(U,C\cup D,\{V_a \mid a\in C\cup D\},\{F_a \mid a\in C\cup D\})$ be a decision Table, in which U is a set of all objects, C is a set of condition attributes, $C=\{C_1,C_2\cdots C_n\}$, D is a set of decision attributes, V_a is value domain of attribute a, F_a is a mapping from set U to V_a. Since that when set D contain some decision attributes, we can always transform it to equivalent form that one decision attribute contain many values, then in this paper we always suppose $D=\{d\}$ for short.

Definition 6 Let $M=\bigcup_{a\in C\cup D} V_a$, then (U,M,I) is a *formal context corresponding to decision Table* S.

In order to illustrates conveniently, for the formal context (U,M,I) in definition 6, the element m of set $\bigcup_{a\in C\cup D} V_a$ is called attributes of formal context (U,M,I), the element m of set $\bigcup_{a\in C} V_a$ called *condition attributes* denoted by V_C and the element m of set $\bigcup_{a\in D} V_a$ called *decision attributes* denoted by V_D. Besides, in this paper we always suppose that if $a,b\in C, a\neq b$, then we have $V_a \bigcap V_b=\phi$.

Definition 7 Let $X\to Y$ be a value dependency that hold in formal context *K* $=(U,M,I)$. If $Y\bigcap V_d\neq\phi$, $X\to Y\bigcap V_d$ is called *decision rules*; if $(Y-X)\bigcap V_C\neq\phi$, $X\to(Y-X)\bigcap V_C$ is called *association rules*;

Definition 8 Let *K*$=(U,M,I)$ be a formal context, $X,Y\subseteq M$, $X\to Y$ is called compact dependencies if it satisfy the following conditions:

(1) for any $X_1\subset X, X_1\to Y$ not hold in *K*;

(2) for any $Y_1\supset Y, X\to Y_1$ not hold in *K*.

Lemma 1 Let *K*$=(U,M,I)$ be a formal context, Y be a intent, $X\subseteq Y$, then $g(X)=g(Y)$ if and only if X is a transversal of hypergraph $\mathrm{H}(g(Y),Y)$.

Theorem 1 [1] $X\to Y$ is compact dependencies if and only if : (1) Y is intent;(2) X is a minimal transversal of hypergraph $\mathrm{H}(g(B),B)$.

Proof. For more detail information, see the reference [1].

Definition 9 Let *K*$=(U,M,I)$ be a formal context, $\sum$ be the set of all compact dependencies that hold in *K*. Then we have:

(1) For all $X\to Y\in\sum$, if $Y\bigcap V_d\neq\phi$, the set of all $X\to Y\bigcap V_d$ is called *decision rules set*, denoted by $\sum_{DR}$;

(2) For all $X\to Y\in\sum$, if $(Y-X)\bigcap V_C\neq\phi$ and $X\bigcap V_d=\phi$, the set of $X\to(Y-X)\bigcap V_C$ is called *association rules set of condition attributes*, denoted by $\sum_{AR}$;

Definition 10 Let $\sum_R$ be a set of decision(*association*) rules. For any rules $X\to Y\in\sum_R$, $\sum_R$ is called *briefest decision(association) rules set* if there not exist $X_1\subset X$ and $X_1\to Y$ hold in the context.

According to definition 10, the decision rules set $\sum_{DR}$ and association rules set $\sum_{AR}$ are not the briefest rules set and need dropped some unsatisfactory rules. For decision rules set $\sum_{DR}$ generated in definition 9 and $X\to Y\in\sum_{DR}$, if there exist $X_1\subset X$ and $X_1\to Y\in\sum_{DR}$, the new decision rules set is a briefest decision rules set after deleting $X\to Y$,i.e. $\sum_{DR}=\sum_{DR}-X\to Y$. The same principles apply to the association rules set.

Theorem 2 All the decision rules hold in context *K* can deduced by adding attributes to front component on a decision rule in the briefest decision rules set $\sum_{DR}$.

Proof. We only need to proof that if decision rule $X \to Y$ （$Y \in V_d$） hold in context *K*, it must exist a decision rule $X_0 \to Y$ such that $X_0 \subseteq X$. Because each decision rule in $\sum_{DR}$ is obtained from a compact dependency $X_0 \to Y_0$ in *K* through $Y_0 \cap V_d$, we only need to proof that if $X \to Y$ hold, it must exist a compact $X_0 \to Y_0$ such that $Y \subseteq Y_0$, $X_0 \subseteq X$.

Let $Y_0 = f(g(X \cup Y))$, it is immediate to see that $Y \subseteq f(g(Y))$ and $f(g(X \cup Y)) \supseteq f(g(Y))$, hence we have $Y \subseteq Y_0$. If decision rule $X \to Y$ hold in *K*, then $g(X) \subseteq g(Y)$. Therefore $g(X) = g(X) \cap g(Y) = g(X \cup Y)$. Because $Y_0 = f(g(X \cup Y))$ is a intent, $f(g(X)) = f(g(X \cup Y)) = Y_0$, then it follows that $g(f(g(X))) = g(Y_0)$. We also show that $g(f(g(X))) = g(X)$, then $g(X) = g(Y_0)$. According to Lemma1, we can show that X is a transversal of $H(g(Y_0, Y_0))$. So it must exist $X_0 \subseteq X$ and X_0 is a minimal transversal of $H(g(Y_0, Y_0))$ such that $X_0 \to Y_0$.

Theorem 3 All the association rules of condition attributes that hold in context *K* can deduced by adding attributes to front component or reducing attributes from back component on a association rule in the briefest association rules set $\sum_{AR}$.

Proof. We only need to proof that if decision rule $X \to Y$ （$Y \subseteq V_C$） hold in context *K*, it must exist a association rule $X_1 \to Y_1$ such that $X_1 \subseteq X$, $Y_1 \supseteq Y$. Because each association rule in $\sum_{DR}$ is obtained from a compact dependency $X_0 \to Y_0$ in K through $X_0 = X_1$, $Y_1 = Y_0 \cap V_C$. We only need to proof that if $X \to Y$ hold, it must exist a compact $X_0 \to Y_0$ such that $Y \subseteq Y_0$, $X_0 \subseteq X$. The detailed proof procedure is similar to the above one.

4. Mining Briefest Rules Based On Compact Dependencies

Two algorithms are proposed in this part. Algorithm 1 can acquire all compact dependencies through a concept lattice corresponding to a formal context. The minimal transversal of each concept (α, β) is computed in Line 8-13 and stored into variable MT. Algorithm 2 can find out all the briefest decision rules and association rules by inputting all the compact dependencies.

Algorithm 1: compute all the compact dependencies

Input: Concept lattice $B(U, M, I)$

Output: The set $\sum_0$ of all the compact dependencies

(1) $\sum_0 = \phi$

(2) For each $(\alpha, \beta) \in B(U, M, I)$

(3) $wv = \phi$

(4) For each $(\alpha_1, \beta_1) \in Parent((\alpha, \beta))$

(5) $wv = wv \cup \{\beta / \beta_1\}$

(6) Next For

(7) ‘compute minimal transversal MT of waned-value hypergraph of concept (α, β)

(8) If $wv = \{W_1, W_2, \cdots, W_k\}$, then

(9) $Tr = \{\{w_1, w_2, \cdots, w_k\} \mid w_i \in W_i, i = 1, 2, \cdots k\}$ ‘Tr is a set of all the transversal

(10) $MT = \phi$
(11) For each $t \in Tr$
(12) If $\forall t^{'} \in Tr, t^{'} \not\subset t$ then $MT = MT \bigcup \{t\}$
(13) Next For
(14) For each $t \in MT$
(15) $\Sigma_0 = \Sigma_0 \bigcup (t \to \beta)$
(16) Next For
(17) Next For
(18) Output $\sum_0$

Algorithm 2: Find out all the briefest decision rules and association rules with confidence of 1.
Input: The set $\sum_0$ of all the compact dependencies in formal context *K*.
Output: all the briefest decision rules and association rules of condition attributes with confidence of 1.
(1) $\sum_{AR} = \phi$, $\sum_{DR} = \phi$
(2) For each $A \to C \in \sum_0$
(3) $B = C - A$
(4) If $B \neq \phi$ and $g(A) \neq \phi$ then
(5) If $B \bigcap V_d \neq \phi$ then
(6) For each $(\varphi \to (d = d_t)) \in \Sigma_{DR}$
(7) If $d_t = B \bigcap V_d$ then
(8) If $\varphi \subseteq A$ then GoTo (14)
(9) If $A \subset \varphi$ then $\Sigma_{DR} = \Sigma_{DR} - (\varphi \to (d = d_t))$
(10) End If
(11) Next For
(12) $\sum_{DR} = \sum_{DR} \bigcup (A \to (B \bigcap V_d))$
(13) End If
(14) If $B \bigcap V_C \neq \phi$ and $A \bigcap V_d = \phi$ then
(15) For each $(\varphi \to \psi) \in \Sigma_{AR}$
(16) If $\psi = B \bigcap V_C$ then
(17) If $\varphi \subseteq A$ then GoTo (24)
(18) If $A \subset \varphi$ then $\Sigma_{AR} = \Sigma_{AR} - (\varphi \to \psi)$
(19) End If
(20) Next For
(21) $\sum_{AR} = \sum_{AR} \bigcup (A \to (B \bigcap V_C))$
(22) End If
(23) End If
(24) Next For
(25) Output $\sum_{AR}$, $\sum_{DR}$

5. Application to Production Data Analysis

Table 1 is a decision Table of production data after discretizing corresponding to part data in blast furnace. In Table1, first three attributes are condition attributes, the last attribute is decision attributes. In order to discover underlying information in production data and improve the ability to make decisions, this paper attempt to find a way to apply the rules extraction algorithm to actual data from a steel works. Table 2 is a formal context converted from the decision Table according to definition 6. Figure 1 is concept lattice corresponding to a formal context in Table 2 and the waned-value is marked at the side of edge between concepts.

Table 1. A Decision Table

a	b	c	d	e
2	1	1	1	2
1	2	1	2	1
1	1	2	2	1
1	1	1	1	2
1	1	2	2	1
2	2	2	2	2

Table 2. The Formal Context Corresponding to the Decision Table

	$a1$	$a2$	$b1$	$b2$	$c1$	$c2$	$d1$	$d2$	$e1$	$e2$
1		×	×		×		×			×
2	×			×	×			×	×	
3	×		×			×		×	×	
4	×		×		×		×			×
5	×		×			×		×	×	
6		×		×		×		×		×

Part of compact dependencies obtained as follows through algorithm 1:
$\{d2 \to d2, a1 \to a1, b1 \to b1, a1d2 \to a1d2e1, e1 \to a1d2e1, c2 \to c2d2\ a1b1 \to a1b1, c1 \to c1, e2 \to e2, b1c2 \to a1b1c2d2e1, b1d2 \to a1b1c2d2e1, a2c2 \to a2b2c2d2e2 \cdots\}$.

For each value dependency $A \to C$ in the above set, compute $A \to (C - A) \bigcap V_d$: $\{a1d2 \to e1, b1c2 \to e1, b1d2 \to e1, a1c2 \to e2, b1c1 \to e2, d1 \to e2, a2 \to e2, a1b2 \to e1, b2c1 \to e1, c1d2 \to e1, a1b1c1 \to e2, a1d1 \to e2, a2b1 \to e2, a2c1 \to e2, a2d1 \to e2, a2b2 \to e2, a2d2 \to e2, a2c2 \to e2\}$.

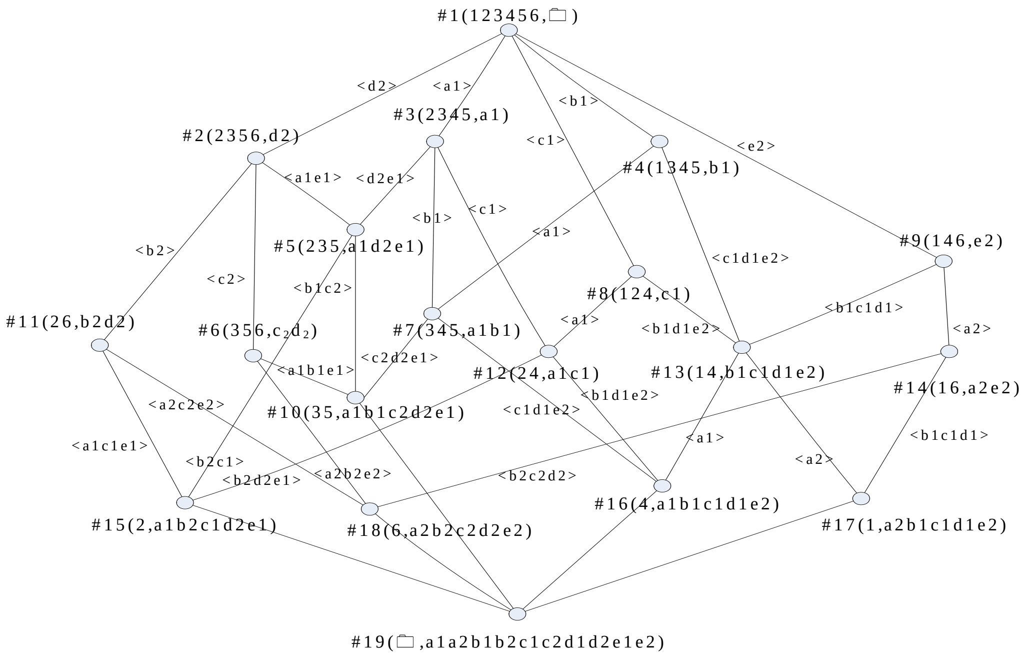

Figure 1. The concept lattice corresponding to formal context

According to algorithm 2, some decision rules have not been inserted to the set of briefest decision rules, satisfying the following condition: their back components are the same as the back component of other decision rules in the set and their front components are super set of the front component of other decision rules. these decision rules include:
$\{a1d1 \to e2, a2d1 \to e2, a2b1 \to e2, a2c1 \to e2, a2b2 \to e2, a2d2 \to e2, \quad a2c2 \to e2, a1b1c1 \to e2\}$

Finally, we acquire all the briefest decision rules: $\{a1d2 \to e1, b1c2 \to e1, b1d2 \to e1, a1c2 \to e2, b1c1 \to e2, d1 \to e2, a2 \to e2, a1b2 \to e1\ , b2c1 \to e1, c1d2 \to e1\}$. Line 14-24 in algorithm 2, for each value dependency $A \to C$ ($A \cap V_d = \phi$), compute $A \to C - A - V_d$ and the result is as follow: $\{b1c2 \to a1d2$, $b1d2 \to a1c2$, $a1c2 \to b1d2$, $b2 \to d2$, $c1b1 \to d1$, $d1 \to b1c1$, $a1b2 \to c1d2$, $c1b2 \to a1d2$, $c1d2 \to a1b2$, $a1b1c1 \to d1$, $a1d1 \to b1c1$, $a2b1 \to c1d1$, $a2c1 \to b1d1$, $a2d1 \to b1c1$, $a2b2 \to c2d2$ $a2d2 \to b2c2$, $b2c2 \to a2d2$, $a2c2 \to b2d2\}$。

Delete some decision rules $\{a1d1 \to b1c1, a2d1 \to b1c1, a1b1c1 \to d1\}$, because there exist some other decision rules in the same set such that their back components are the same, but their front components are super set of the latter. All the briefest association rules of condition attributes are as follows:$\{b1c2 \to a1d2$, $b1d2 \to a1c2$, $a1c2 \to b1d2$, $b2 \to d2$, $c1b1 \to d1$, $d1 \to b1c1$, $a1b2 \to c1d2$, $c1b2 \to a1d2$, $c1d2 \to a1b2$, $a2b1 \to c1d1$, $a2c1 \to b1d1$, $a2b2 \to c2d2$ $a2d2 \to b2c2$, $b2c2 \to a2d2$, $a2c2 \to b2d2\}$。

Through the example, we can analysis the relationship between many process parameters and aim parameters,as well as process parameters and other process parameters.

6. Conclusion

In this paper, we have proposed an approach for briefest rules extraction based on the theory of compact dependencies. One of the main features of our approach is their generic feature and they can apply in several application fields. Our aim when proposing these methods is to make use of concept lattice to solve database and data mining problems. This makes it possible to take advantage of existing and efficient concept lattice tools. The final objective is to achieve a unified software platform devoted to enterprise database decision-making. So this text has very important realistic meanings to the research and discussion of the characteristic and rules mining tactics of concept lattice.

References

[1] MA Yuan, ZHANG Xue-Dong, CHI Cheng-Ying. Compact Dependencies and Intent Waned Values. *Journal of Software.* 2011; 22(5): 962-971.

[2] Boris Milovic. Prediction and decision making in Health Care using Data Mining. *International Journal of Public Health Science (IJPHS)*. 2012; 1(2): 69-78.

[3] Chandrashekhar AM, Raghuveer K. Performance evaluation of data clustering techniques using KDD Cup99 Intrusion detection data set. *International Journal of Information and Network Security (IJINS)*. 2012; 1(4): 294-305.

[4] Bertossi L, Hunter A, Schaub T. *Editors*. On the computational complexity of minimal change integrity maintenance in relational databases, Inconsistency Tolerance. LNCS, Springer, Heidelberg. 2005; 3300: 119-150.

[5] T Eiter, G Gottlob. Identifying The Minimal Transversals of A Hypergraph and Related Problems. *SIAM Journal on Computing.* 1995; 24(6): 1278-1304.

[6] T Eiter, G Gottlob. Hypergraph Transversal Computation and Related Problems in Logic and AI. *In European Conference on Logics in Artificial Intelligence (JELIA'02)*. 2002: 549-564.

[7] Guigues J, Duquenne V. Families Minimales d'implications Informatives Resultants d'un Tableau de Donnees Binaires. *Mathematics and Social Sciences.* 1986: 495-518.

[8] Zaki MJ. Mining non-redundant association rules. In: Data Mining and Knowledge Discovery (DMKD), *Springer*, Heidelberg. 2004; 9(3): 223–248.

[9] Ho TB, Cheung D, Liu H. *Editors*. A new informative generic base of association rules. PAKDD LNCS (LNAI), Springer, Heidelberg. 2005; 3518: 81–90.

[10] Li J. On optimal rule discovery. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*. 2006; 18(4): 460–471.

[11] Bertet K, Monjardet B. The multiple facets of the canonical direct unit implicational basis. Theoretical Computer Science 411. 2010; 22-24: 2155-2166.

[12] T Hamrouni, S Ben Yahia, E Mephu Nguifo. *Generic association rule bases: are they so succinct? CLA'06*. Proceedings of the 4th international conference on Concept lattices and their applications. 2006: 198-213.

[13] Ganter B, Wille R. Formal Concept Analysis: Mathematical Foundations. *Springer.* Berlin. 1999.