文章编号:1001 - 9081(2010)04 - 1072 - 04

多维概念格与关联规则发现

郭显娥[1],王俊红[2]

(1. 山西大同大学 数学与计算机科学学院,山西 大同 037009; 2. 山西大学 计算机与信息技术学院,太原 030006)

(scjguo@ yahoo. com. cn)

摘　要:在引用多维数据序列对概念内涵进行不同维度的描述的基础上,提出了多维概念格的形式化定义及其构造方法;并给出了基于多维概念格的关联规则提取方法,该方法通过发现最大频繁多维数据序列研究不同维度属性之间的依赖关系。实例表明,多维概念格利于发现内容更丰富的有用信息。

关键词:概念格;多维概念格;维度;频繁多维序列;多维精简序列

中图分类号: TP18;TP311. 13　　　　**文献标志码**:A

Multi-dementional concept lattice and association rules discoverey

GUO Xian-e[1], WANG Jun-hong[2]

(1. *School of Mathematics and Computer Science, Shanxi Datong University, Datong Shanxi* 037009, *China*;
2. *School of Computer and Information Technology, Shanxi University, Taiyuan Shanxi* 030006, *China*)

Abstract: Based on the description of deferent dimension of concept intent using multi-dimensional data sequence, formal definition and construction method of multi-dimensional concept lattice were proposed in the paper. Also the association rules discovery method based on multi-dimensional concept latticewas given, which studied dependence relation among deferent dimension attributes thought discovering biggest frequent multi-dimensional data sequence. Examples show that multi-dimensional concept lattice helps to discover richer useful information.

Key words: concept lattice; multi-dementional concept lattice; dimensionality; frequent multi-dementional sequence; multi-dementional tidy sequence

0　引言

由外延和内涵两个部分所组成的单元称为概念,基于对概念的理解,文献[1]中提出了形式概念分析,用于概念的发现、排序和显示。形式概念分析及概念格(Galois 格)模型已经被广泛地应用于多个领域。

在数据库知识发现(Knowledge Discovery in Database, KDD)中,关联规则是很有价值的一类规律,它指的是一个形如 $X \Rightarrow Y$ 的表达式,其中 X 和 Y 是特征集合,其直观含义是:数据库中具有特征 X 的行(或记录、对象)可能也具备特征 Y。自从文献[2]中提出从大型数据库中挖掘关联规则以来,关联规则发现已经得到了广泛而深入的研究,并成为 KDD 的核心任务之一。文献[3]中提出了由概念格提取蕴涵规则的算法,文献[4]提出了近似蕴涵规则的提取算法,文献[5 - 6]以概念格为基础,分析了概念格和关联规则发现之间的关系,并给出了相应格结构的渐进式算法和关联规则提取算法。

对形式概念分析及概念格研究较多的是概念的内涵与外延的性质,但对概念产生的背景研究得较少。文献[7]以解决多维序列模式挖掘算法为出发点,提出了多维概念格的定义,将概念的内涵分为基础维与任务维,该模型是建立在概念格基础上的,并没有突破传统概念内涵,没有建构真正意义的空间维度概念。

本文充分研究了概念产生的背景,从空间维度出发,引用多维数据序列对内涵进行不同维度的描述,提出多维概念格的形式化定义,并构造多维概念格格结构模型,利于发现内容更丰富的有用信息。而基于多维概念格的关联规则提取算法是建立在概念格与多维数据序列挖掘基础上,通过发现最大频繁多维数据序列,研究不同维度中属性的依赖关系。

1　多维概念格的形式化定义

1.1　相关概念

多维概念格是基于多维数据序列建立的数据结构,下面引用文献[7 - 8]对多维数据序列进行形式化描述,并在此基础上给出多维概念格的形式化定义。

定义 1[7]　n 维数据序列 A 定义为 $n-1$ 维数据序列构成的有序列表,记为 $A = \langle a_1 a_2 \cdots a_k \rangle_n$,其中 $n(n \geqslant 1)$ 为维数,即维度的个数,a_i 为 $n-1$ 维序列。当 $n=1$ 时,a_i 称为项元素或项,记第 k 个项为 $A(k)$,其中 $1 \leqslant k \leqslant card(A)$。

依据定义 1,n 维数据序列是由 $n-1$ 维序列构成的。例如一个三维数据序列 A,设 $A = \langle\langle\langle ab\rangle_1\langle cd\rangle_1\rangle_2\langle\langle bd\rangle_1\langle bc\rangle_1\rangle_2\rangle_3$。$A$ 是由两个二维序列 $\langle\langle ab\rangle_1\langle cd\rangle_1\rangle_2$ 与 $\langle\langle bd\rangle_1\langle bc\rangle_1\rangle_2$ 构成,第一个二维序列由 $\langle ab\rangle_1$ 与 $\langle cd\rangle_1$ 构成,构成一维序列的元素 a,b,c,d,b,d,b,c 称为项元素或称为项。这里的下标号1,2,3 分别代表3 个不同

收稿日期:2009 - 11 - 02;**修回日期**:2009 - 12 - 14。　　**基金项目**:天津市自然科学重大基金资助项目(07JCZDJC06500)。

作者简介:郭显娥(1964 -),女,山西浑源人,副教授,硕士,主要研究方向:概念格、知识发现;　王俊红(1979 -),女,山西曲沃人,讲师,博士,主要研究方向:概念格、粗糙集。

的维度。

定义 1 所描述的 n 维数据序列形式虽然容易理解,但是相对复杂,算法实现比较困难。为此引入 n 维数据序列的压缩形式。

定义 2[7]　给定 n 维数据序列 A , $A(k)$ 与 $A(j)$ 为第 k 与第 j 个项元素,如果存在 i 维序列 B 使得 $A(k) \in B$, $A(j) \in B$,且不存在满足上述条件的 u ($u < i$) 维序列,则称项元素 $A(k)$ 与 $A(j)$ 的维层跨度(Dimansional Span,DS)为 i ,记为 $DS(A,k,j) = i$ 。特别地,如果 $A(k)$ 与 $A(j)$ 隶属于同一个项集(由项元素组成的集合),则 $DS(A,k,j) = 0$ 。

在 $A = \langle\langle\langle ab\rangle_1\langle cd\rangle_1\rangle_2\langle\langle bd\rangle_1\langle bc\rangle_1\rangle_2\rangle_3$ 中, $DS(A,1,2) = 1$,表示数据序列A的第1与第2项可以在一维空间上并存, $DS(A,2,3) = 2$,则表示第2、3项至少需要在二维空间上才能并存, $DS(A,4,5) = 3$,依此类推。为便于更好地理解定义 1 和定义 2,我们将上例中的序列 A 以树的形式进行表示(如图1所示),其中根节点代表 A 的三维序列,第一层节点代表A的二维序列,第二层节点代表 A 的一维序列,终节点是项元素。对任意两个项 $A(k)$ 与 $A(j)$,包含它们的最近的共同祖先就是定义 2 中的序列 B 。如果 B 是 A 的 i 维成员序列,则有 $DS(A,k,j) = i$ 。

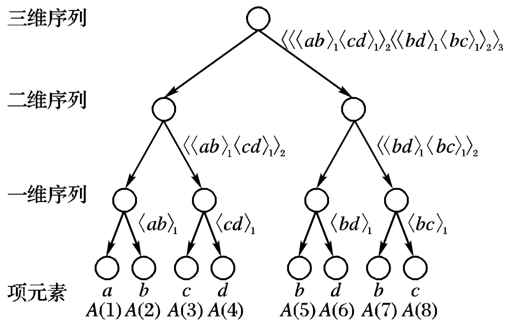

图1　三维数据序列(A) 的树结构

定义 3[7]　任何 n 维数据序列 A 可以等价变换为压缩形式 $\langle a_{i_1}a_{i_2}\cdots a_{i_p}\rangle_n$, 其中 a_{i_k} 由两部分构成: 一部分是项 $a_{i_k}.item = A(i)$, 另一部分是该项的维层跨度, $a_{i_k}.span = DS(A,i-1,i)$,特殊地规定: $a_{i_1}.span = 1$ 。

根据定义 3, 三维数据序列 $A = \langle\langle\langle ab\rangle_1\langle cd\rangle_1\rangle_2\langle\langle bd\rangle_1\langle bc\rangle_1\rangle_2\rangle_3$ 的压缩形式为 $A = \langle a_1b_1c_2d_1b_3d_1b_2c_1\rangle_3$ 。

在压缩形式下,项及维层跨度是构成数据序列的重要元素。

推论 1　压缩形式下两个维数相同的数据序列相等的充分必要条件是:项元素对应相同且维层跨度一样。

下面给出压缩形式下两个维数相同的数据序列包含与交运算的定义。

定义 4[7]　设 A,B 是以压缩形式表示的两个 n 维数据序列,分别表示为 $\langle a_1a_2\cdots a_p\rangle_n$, $\langle b_1b_2\cdots b_q\rangle_n$,如果存在整数 $i_1 < i_2 < \cdots i_r (1 \leqslant i_k \leqslant p, p = card(A)$ 满足:

1) $b_k.item = a_{i_k}.item$

2) $b_k.span = DS(A,i_{k-1},i_k)$　$(1 \leqslant k \leqslant q)$

则称 B 为 A 的子序列,或称 A 包含 B,记为 $B \subset A$。

例如, $\langle a_1b_1b_3c_2\rangle_3 \subset \langle a_1b_1c_2d_1b_3d_1b_2c_1\rangle_3$,而 $\langle a_1b_1b_3c_1\rangle_3$ 则不是 $\langle a_1b_1c_2d_1b_3d_1b_2c_1\rangle_3$ 的子序列。

定义 5[7]　两个 n 维数据序列的交集定义为其所有最大公共子序列所构成的集合。设 A、B 分别为两个 n 维数据序列,则 A 和 B 的交集 $A \cap B = \{C \mid C \subseteq A \text{ and } C \subseteq B\}$,且不存在 $C' \supset C$,满足 $C' \subseteq A$ and $C' \subseteq B$。

例 1　由多维序列 $A = \langle a_1b_1c_2d_1a_3c_1\rangle_3$ 与 $B = \langle a_1b_1b_3d_1b_3c_1\rangle_3$ 构成的集合记为 M_n, 即 $M_n = \{\langle a_1b_1c_2d_1a_3c_1\rangle_3, \langle a_1b_1b_3d_1b_3c_1\rangle_3\}$,则有如下结论: $\langle a_1b_1\rangle_3$ 分别为 A 与 B 的子序列; $A \cap B = \{\langle a_1b_1c_3\rangle_3, \langle d_1c_3\rangle_3\}$ 。

1.2　多维概念格定义

根据概念格的基本定义和多维数据序列的特点,给出多维概念格的定义。

定义 6　称 $K = (G, M_n, I)$ 为多维序列形式背景,其中 G 为有限非空事例集合(称为外延), M_n 为压缩形式的 n 维数据序列集合(称为内涵), I 是 G 与 M_n 之间的一个二元关系,即 $I \subseteq G \times M_n$:对于 xIy_n ,表示 x 与 y_n 间(其中 $x \in X, y_n \subset Y, X \subseteq P(G), Y \subseteq M_n$) 存在关系 I,即事例 x 拥有属性构成的数据序列 y_n。则在此形式背景下,存在唯一的偏序集与之相对应,并且这个偏序集产生唯一的格结构。格中的每一个元素是一个二元组(X,Y)。

定义 7　当且仅当二元组(X,Y) 满足性质:

$X = g(Y)$: $g(Y) = \{x \in G \mid y_n \subset Y, \ xIy_n\}$

$Y = f(X)$: $f(X) = \{y_n \subset M_n \mid x \in X, \ xIy_n\}$

时,称二元组(X,Y) 关于 I 是完备的。

定义 8　由定义6和定义7所诱导的格L就称为多维概念格。它与概念格不同之处在于,概念的内涵是 n 维有序的数据序列,相当于属性串集合。

推论 2　多维概念格 L 一定是概念格。

显然,对于形式背景 $K = (G, M_n, I)$,定义关系"$\leqslant$" 为"$\subseteq$",则它是集合 L 上的一个偏序(满足自反性、反对称性和传递性),由此可诱导出 L 上的一个格结构。可以证明,它是一个完备格。相应的下确界和上确界定义为:

$$\bigwedge_{t \in T}(X_t, Y_t) = \left(\bigcap_{t \in T} X_t, f\left(g\left(\bigcup_{t \in T} Y_t\right)\right)\right)$$

$$\bigvee_{t \in T}(X_t, Y_t) = \left(g\left(g\left(\bigcup_{t \in T} X_t\right)\right), \bigcap_{t \in T} X_t\right)$$

其中$(X, Y) \in L$, T 是指标集,此完备格即为形式背景 K 的概念格。

定理 1　多维概念格中的所有节点都是最大扩展序偶。

格中每个节点(X,Y,I) 是一个序偶,且对于关系 I 都是完备的,显然可见所有节点都是最大扩展序偶,即只有最大扩展的序偶才出现在多维概念格的层次结构中。

定理 2　在多维概念格节点 $C1 = (X1, Y1)$ 和 $C2 = (X2, Y2)$,若 $X1 \subseteq X2$ (等价于 $Y1 \supseteq Y2$),则有 $C1 \leqslant C2$。若 $C1 \leqslant C2$,且不存在元素 $C3$,使得 $C1 < C3 < C2$,则存在边 $C1 \leftarrow C2$。称 $C1$ 为 $C2$ 的子节点, $C2$ 为 $C1$ 的父节点。

2 多维概念格格结构的构造

2.1 相关概念

定义9 给定形式背景 $K=(G,M_n,I)$，有多维序列 $P(P\subseteq M_n)$，数据集中包含 P 的记录数与数据集中总记录数的比值称为 P 的支持度，记为 $support(P)$。

定义10 给定多维序列 $P(P\subseteq M_n)$ 的最小支持度阈值 $min\_support$，如果 $support(P)\geqslant min\_support$，则称 P 为频繁多维序列。

例2 设有形式背景 $K=(G,M_n,I)$，其中 G 为事务集合 $G=\{1,2,3\}$，M_n 为事务具有的属性串，$M_n=\{\langle b_1c_1d_2e_1b_3a_1\rangle_3,\langle b_1d_1c_3b_1a_2\rangle_3,\langle a_1b_1a_3d_1\rangle_3\}$，$I$ 描述 G 中元素拥有 M_n 中的序列。

表1 G 与 M_n 关系

G	M_n
1	$\langle b_1c_1d_2e_1b_3a_1\rangle_3$
2	$\langle b_1d_1c_3b_1a_2\rangle_3$
3	$\langle a_1b_1a_3d_1\rangle_3$

假定多维序列的最小支持度为50%，则对 $P=\langle b_1b_3\rangle_3$，其中有两条记录(第1和2条记录)包含 P，即 $support(P)=2/3$，因此 P 是一个频繁多维序列。

定理3 频繁多维序列 P 的任一子序列都是频繁序列。

证明 设 N 为数据集中记录总数。假设 P 是 m 个记录子集，若 $P'\subset P$，则 P' 是 $m'(m'>m)$ 个记录子集，因此 $m/N\geqslant min\_support$，$m'/N\geqslant min\_support$ 也成立。

显然，频繁多维序列 P 中的项一定是频繁项。 证毕。

定义11 不包含非频繁项的多维序列称为多维精简序列。

2.2 多维概念格构造思想

基于多维概念格内涵的特殊性，建格前首先要对构成内涵的多维数据序列进行处理，然后按照渐近算法构造格节点，大致分为以下两个步骤进行。

第一步 内涵数据预处理。

1) 将内涵中的多维数据序列转变为多维压缩序列(定义3)；

2) 扫描数据集，删除非频繁项；

3) 由所有的频繁项，得到多维精简序列。

由压缩序列到精简序列，需要删除非频繁项，下面给出删除项的定义。

定义12[8] 设 $A=\langle a_1a_2\cdots a_i\cdots a_p\rangle_n$ 为多维压缩序列，从 A 中删除项 a_i 后变为 $A'=\langle a_1a_2\cdots a_{i-1}a_{i+1}\cdots a_p\rangle_n(2\leqslant i\leqslant p-1$，$p$ 为 A 的基数)，a_{i+1} 维层跨度值要发生变化，定义 $a_{i+1}.span=\max\{a_i.span,a_{i+1}.span\}$，当 $i=1$ 时，$a_{i+1}.span=1$。

例3 设有形式背景 $K=(G,M_n,I)$，其中 $G=\{1,2,3\}$：

$$M_n=\{\langle\langle\langle ab\rangle_1\langle cd\rangle_1\rangle_2\langle\langle ac\rangle_1\rangle_2\rangle_3,\ \langle\langle\langle ab\rangle_1\rangle_2\langle\langle fd\rangle_1\rangle_2\langle\langle bc\rangle_1\rangle_2\rangle_3,\ \langle\langle\langle bc\rangle_1\langle ac\rangle_1\rangle_2\langle\langle h\rangle_1\langle ab\rangle_1\rangle_2\rangle_3\}$$

I 描述 G 中元素拥有 M_n 中的多维数据序列。设最小支持度 $min\_support=2$，列出 M_n 数据预处理过程。

解 因为 $min\_support=2$，故 a,b,c,d 为频繁项，而 f,h 为非频繁项，因此需要删除包含 f,h 的内容。

1) M_n 转变为多维压缩序列：

$$M_n=\{\langle a_1b_1c_2d_1a_3c_1\rangle_3,\langle a_1b_1f_3d_1b_3c_1\rangle_3,\ \langle b_1c_1a_2c_1h_3a_2b_1\rangle_3\}$$

2) M_n 转变为多维精简序列：

$$M_n=\{\langle a_1b_1c_2d_1a_3c_1\rangle_3,\langle a_1b_1d_3b_3c_1\rangle_3,\langle b_1c_1a_2c_1a_3b_1\rangle_3\}$$

第二步 内涵预处理为多维精简序列后，建立多维概念格渐进算法。

多维概念格渐进式构造的基本思想是，当渐进地追加一个新对象时，要依据概念内涵的数据序列变化情况更新原概念格，同时修改相应的边，最后形成新的多维概念格。

给定一个多维概念格 L，其中任一多维概念 $C=(G,A)$，新追加的概念 $O=(x,Y)$，其中 Y 为多维精简序列。令更新后的多维概念格 L'，新的多维概念格 L' 与原多维概念格 L 有如下三种关系。

1) 不变格，即 L' 就是 L。此时 $Y\cap A=\varnothing$，即新增概念内涵与原概念内涵没有共同子序列。

2) 更新格，即 L' 保持 L 的内涵数据序列，外延扩展。此时 $A\subset Y$，即新增概念内涵序列包含原概念内涵序列(包含关系的判断依据定义4)，得到更新节点为 $C'=(G\cup\{x\},A)$。

3) 新增概念，若不满足条件(1)、(2)，则 L' 由 L 新增概念所得。此时$(A\cap Y)=B\neq\varnothing$，新增节点为 $N=(G\cup\{x\},B)$。

对于新增节点，找到其父节点、子节点，分别加一条新增节点到父节点的边。若在父、子节点之间存在一条边，则直接删除。

2.3 多维概念格应用

多维概念格应用于空间数据背景，格中节点内涵的描述具有明确的维度标志，可以应用于各个领域，如：

1)用于零售数据挖掘中，识别顾客购买行为，发现顾客在不同时间、地点购买模式和趋势，得出顾客的消费习惯，据此对商品的价格和促销手段进行调整；

2)用于DNA序列模式研究，可以从中找出导致各种疾病的特定的基因序列模式；

3)用于Web用户数据模式的挖掘，通过分析Web日志可以使网管员改善网站的页面组织，针对特定用户进行Web页面存取，对个体用户定制Web服务，尽量为大多数访问者的浏览提供方便；

4)应用于客户关系管理(CRM)、医疗诊断、自然灾害预测等。

例4给出一个简单的应用实例。

例4 表2是某班学生选课情况表，G 为学生序号(共3人)，学生所选课程名用字母表示，I 为 G 中的对象拥有的多维数据序列。该例中课程名、必修课程与选修课程、春季与秋季，分别对应一、二、三3个维度，是一个三维形式背景。

形式背景：$K=(G,M_n,I)$ 中 $G=\{1,2,3\}$，M_n 是一个三维序列，其精简序列为：

$$M_n = \{\langle a_1b_0c_0D_2E_0x_3y_0W_2Y_0Z_0\rangle_3, \langle a_1d_0C_2F_0G_0z_3W_2X_0\rangle_3, \langle e_1f_0D_2G_0x_3z_0X_2Y_0\rangle_3\}$$

$min\_support = 2$

得到相应的 Hasse 图，如图 2 所示。

表 2　学生选课情况表

I	春季		秋季	
	必修课	选修课	必修课	选修课
1	*abc*	*DE*	*xy*	*WYZ*
2	*ad*	*CFG*	*z*	*WX*
3	*ef*	*DG*	*xz*	*XY*

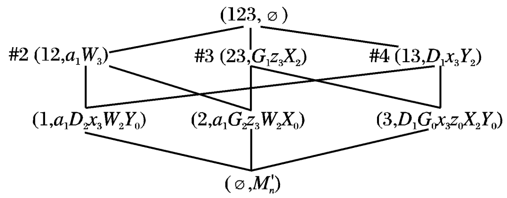

图 2　表 1 对应的 Hasse 图

图 2 中，M_n' 为 M_n 的多维精简序列，即：$M_n' = \{\langle a_1D_2x_3W_2Y_0\rangle_3, \langle a_1G_2z_3W_2X_0\rangle_3, \langle D_1G_0x_3z_0X_2Y_0\rangle_3\}$。

3　基于多维概念格的关联规则提取

在实际应用中最大频繁序列模式往往蕴含的知识更为丰富，应用价值更高。基于多维概念格的完备特性，利于关联规则提取。

3.1　多维概念格与关联规则提取

定义 13[8]　设多维序列 M_n，支持度为 C，如果不存在多维序列 M_n'（支持度为 C'）满足如下条件：

1）$C' \geqslant C$；

2）$M_n' \supset M_n$ 则称 M_n 为最大多维序列。

定理 4　多维概念格中的节点，若内涵序列满足定义 13，则必为最大频繁多维序列。

证明　因为多维概念格中的节点内涵是频繁序列，且满足定义 13，故节点内涵为最大频繁多维序列。　证毕。

3.2　关联规则提取算法

基于多维概念格的关联规则挖掘算法，步骤如下：

1）预处理原始数据集，得到多维序列形式背景 $K = (G, M_n, I)$，将内涵由多维数据序列化简为多维精简序列；

2）构建多维概念格；

3）因为格内节点内涵为最大频繁多维序列，故可直接从格上诱导出关联规则；

4）对于频繁节点 (X, Y)，所有形式的 $A \Rightarrow B$ 都是蕴涵规则，其中，$A, B \subset Y$，$B = Y - A$，且 $(X, Y) \leqslant (extention(A), A)$；

5）生成置信度 $|extention(Y)| / |extention(A)| > min\_conf$ 的关联规则 $A \Rightarrow B$。

例 5　分析例 4 中所蕴含的关联规则。

解　例 4 中所对应的 Hasse 图（如图 2 所示），设用户给定的 $min\_support = 2$ 与 min_conf 为 60%。

根据多维概念格建格特点知，格中节点蕴含其关联规则。如图 2 中 #2，#3，#4 节点内涵即为频繁多维序列，其节点为频繁节点。

由 #2 节点导出规则：$a_1 \Rightarrow W_3$，蕴含的信息是：春季必修 a 课程的学生，秋季将选修 W 课程。

由 #3 节点导出规则：$G_1 \Rightarrow z_3X_2$，蕴含的信息是：春季选修 G 课程的学生，秋季必修 z 课程同时选修 X 课程，等等。

同样的算法作用在不同数据背景的概念格中，提取的规则是不一样的。多维概念格所提取的规则不仅描述概念格属性性质之间的关系，同时也给出各属性背景之间与背景与属性之间的关系。下面列出传统概念格与多维概念格所含信息量比较表。

表 3　传统概念格与多维概念格所含信息量比较

概念格	规则集举例	是否带有维度标志	所含信息量
传统概念格	$G \Rightarrow zX$	否	少，只涉及到所选课程，即选 G 课的必选 z 和 X 课程
多维概念格	$G_1 \Rightarrow z_3X_2$	是	多，涉及到所选课程，课程选修性质以及开课学期

4　结语

本文引入了多维概念格的定义，构造了多维概念格结构，通过引入多维数据序列对内涵进行不同维度的描述，进一步丰富了信息内涵，格结构较为复杂，但能处理的问题也比较广泛。基于多维概念格的关联规则发现，利于找到最大频繁多维数据序列，挖掘蕴含的有用背景信息。多维概念格是概念格的扩展模型，内涵包含了大量丰富的背景信息，进一步发掘和拓展该模型的应用领域，以及利用多维概念格进行知识发现，有待于我们更深入的研究。

参考文献：

[1]　WILLE R. Restructuring lattice theory: An approach based on hierarchies of concepts[C]// Ordered Sets. Dordrecht: Reidel, 1982: 445－470.

[2]　AGRAWL R, IMIELINSKI T, SWAMI A. Mining association rules between sets of items in large databases [C]// Proceedings of 1993 ACM SIGMOD International Conference on Management of Data. New York: ACM, 1993: 207－216.

[3]　GODIN R, MISSAOUI R. An incremental concept formation approach for learning from databases[J]. Theoretical Computer Science, 1994, 133(2):387－419.

[4]　MISSAOUI R, GODIN R. Extracting exact and approximate rules from databases[C]// Proceedings of the SOFTEKS Workshop on Incompleteness and Uncertainty in Information Systems. London: Springer-Verlag, 1994: 209－222.

[5]　谢志鹏，刘宗田. 概念格与关联规则发现[J]. 计算机研究与发展，2000, 37(12):1415－1421.

[6]　梁吉业，王俊红. 基于概念格的规则产生集挖掘算法[J]. 计算机研究与发展，2004, 41(8):1339－1344.

[7]　金阳，左万利. 多维概念格与多维序列模式的增量挖掘[J]. 计算机研究与发展，2007, 44(11):1816－1824.

[8]　金阳，左万利. 有序概念格与 WWW 用户访问模式的增量挖掘[J]. 计算机研究与发展，2003, 40(5):675－683.