

基于属性偏序结构图的关联规则提取方法

王秋婷*, 王诗慧

辽宁师范大学, 数学学院, 辽宁 大连
Email: *a1014wangqt@163.com

收稿日期: 2021年3月20日; 录用日期: 2021年4月21日; 发布日期: 2021年4月28日

摘 要

针对Apriori算法生成大量冗余关联规则的问题, 本文提出了一种基于属性偏序结构图的关联规则提取方法。该方法旨在寻找相同支持度下的最大频繁项目集, 进而提取无冗余关联规则。本文提出的方法不仅减少了挖掘频繁项目集的数量, 从而提高关联规则提取的效率, 而且将关联规则转换成属性偏序结构图中的知识表示形式, 实现了频繁项分层的关联规则可视化展示。具有较强的可读性, 有助于用户对关联规则进行深入分析, 提高对潜在知识的利用和发掘程度。

关键词

Apriori算法, 属性偏序结构图, 无冗余关联规则, 可视化

An Association Rule Extraction Method Based on Attribute Partial Order Structure Diagrams

Qiuting Wang*, Shihui Wang

School of Mathematics, Liaoning Normal University, Dalian Liaoning
Email: *a1014wangqt@163.com

Received: Mar. 20th, 2021; accepted: Apr. 21st, 2021; published: Apr. 28th, 2021

Abstract

Aiming at the problem of Apriori algorithm generating a large number of redundant association rules, this paper proposes an association rule extraction method based on attribute partial order

*通讯作者。

structure diagrams. It can extract non-redundant association rules by finding the maximum frequent item sets under the same support. The method proposed in this paper can reduce the number of mining frequent item sets, so as to improve the efficiency of extracting association rules. Moreover, the association rules are converted into the knowledge representation in the attribute partial order structure diagrams, which realizes the visualization of the association rules of the frequent item hierarchy. With strong readability, it is helpful for users to conduct in-depth analysis of association rules and improve the utilization and exploitation of potential knowledge.

Keywords

Apriori Algorithm, Attribute Partial Order Structure Diagrams, Non-Redundant Association Rules, Visualization

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

关联规则挖掘是数据挖掘研究中的一个重要部分,旨在挖掘数据库中有意义的关联。1993年由Agrawal等[1]提出的Apriori算法,是经典的挖掘布尔型频繁项目集的算法并得到广泛使用。利用Apriori算法获得的关联规则往往存在大量的冗余,而且对规则的可视化存在不足。关联分析与规则的可视化结合,有利于人们发现隐藏的知识,已经引起了许多学者的关注。概念格是一种从形式背景进行数据分析和规则提取的强有力工具,通过生成Hasse图能够直观表示出概念之间的层次关系,并且能够体现概念之间的泛化和例化关系[2]。因此在形式概念分析理论下,基于概念格的关联规则提取的研究倍受学者关注[3][4]。目前,已有很多基于概念格的规则提取算法。胡可云[5]等提出一个从概念格上提取关联规则和分类规则的算法,实现了关联规则和分类规则的挖掘在概念格框架下的统一。王德兴等[6]利用量化概念格获取频繁项目集,方法直观,表示简洁。杨葛英等[7]提出了以概念格为背景的关联规则可视化方法,在概念格中展现关联规则路径。

形式概念分析强调的是属性与对象之间严谨的充要关系,与形式概念分析理论不同,关联规则注重的是属性间联系的紧密程度,而属性偏序结构图恰好提供了较为松散的属性关联研究视角。属性偏序结构图[8]是从形式概念分析理论发展而来的不同原理的知识发现数学方法,它从认知事物的角度出发,挖掘形式背景中对象之间、属性之间以及属性与对象之间的关系。属性偏序结构图中边与边之间不交叉,当形式背景中对象和属性增加时,概念格的复杂程度增大,可读性降低,而属性偏序结构图仍然层次关系明确,因此有助于更直观地进行知识模式发现和规则提取。目前属性偏序结构图在中医药数据挖掘、模式识别分类等相关研究中取得了很好的应用效果[9][10][11]。本文主要研究无冗余关联规则提取方法和关联规则可视化表现形式,提出了一种基于属性偏序结构图的关联规则提取方法。将无冗余关联规则提取问题转化成挖掘数据集中相同支持度下的最大频繁项目集问题,并把频繁项转化成属性偏序结构图中的知识表示,为关联规则的提取提供了新的发现视角。理论分析证明,本文提出的方法是有效的。

2. 预备知识

定义1 [1]称 $K = (U, A, I)$ 为一个形式背景,其中 U 为对象集, A 为属性集, I 为 U 和 A 之间的二元

关系, $I \subseteq U \times A$ 。若 $x \in U, a \in A, (x, a) \in I$, 则说 x 具有属性 a , 记为 xIa 。

定义 2 [1] 在形式背景 $K = (U, A, I)$ 中, 对任意的 $G \subseteq U, B \subseteq A$, 定义

$$f(G) = \{a \in A \mid \forall x \in G, (x, a) \in I\}$$

$$g(B) = \{x \in U \mid \forall a \in B, (x, a) \in I\}$$

$f(X)$ 表示 G 中所有对象共同拥有的属性组成的集合, $g(B)$ 表示 B 中所有属性的对象组成的集合。

定义 3 [12] 在形式背景 $K = (U, A, I)$ 中, 如果属性 $m \in A$ 满足

$$g(m) = U$$

则称属性 m 为形式背景 K 的最大共有属性。

定义 4 [12] 在形式背景 $K = (U, A, I)$ 中, 如果属性 $m_0, m_1, m_2, \dots, m_k \in A$ 满足

$$g(m_t) \subseteq g(m_0), t = 1, 2, 3, \dots, k, k \geq 2$$

则称在形式背景 K 中, 属性 m_0 为属性集合 $\{m_1, m_2, \dots, m_k\}$ 的共有属性。

定义 5 [12] 在形式背景 $K = (U, A, I)$ 中, 如果属性 $m \in A$ 满足

$$|g(m)| = 1$$

则称属性 m 为形式背景 K 的独有属性, 其中 $|g(m)|$ 表示具有集合 $g(m)$ 的基数。

属性偏序结构图的构建步骤[8]。

Step 1: 计算最大共有属性 M_0 及其结点 C_0 , 若不存在最大共有属性, 则增加 (\emptyset, U) 作为顶层结点;

Step 2: 在最大共有属性 M_0 下计算所有共有属性集合 $\{m_{11}, m_{12}, \dots, m_{1p}\}$ 和独有属性集合 $\{s_{11}, s_{12}, \dots, s_{1p}\}$ 及其对应二元组 $\{C_{m1}, C_{m2}, \dots, C_{mp}\}$ 和 $\{C_{s1}, C_{s2}, \dots, C_{sp}\}$, 并与最大共有属性的结点 C_0 间绘制有向边。

如果该层有序对结点与最大共有属性的结点间满足关系:

$$\bigcup_{i=1}^p g(m_{1i}) \cup \bigcup_{i=1}^q g(s_{1i}) = U,$$

则进入下一步; 不满足上式时, 在最底层结点与最大共有属性的结点间存在 k 条有向边。其中,

$$k = \left| U - \bigcup_{i=1}^p g(m_{1i}) \cup \bigcup_{i=1}^q g(s_{1i}) \right|。$$

Step 3: 在上层每个共有属性结点 $\{C_{m1}, C_{m2}, \dots, C_{mp}\}$ 下, 计算上层共有属性 $m_i, t \in (1:p)$ (t 表示上层的层数) 的共有属性 $\{m_{t+1,1}, m_{t+1,2}, \dots, m_{t+1,l}\}$ 和独有属性 $\{s_{t+1,1}, s_{t+1,2}, \dots, s_{t+1,k}\}$ 及其对应二元组结点 $\{C_{t+1,1}, C_{t+1,2}, \dots, C_{t+1,l}\}$ 和 $\{C_{st+1,1}, C_{st+1,2}, \dots, C_{st+1,k}\}$, 在每个独有属性的二元组结点 $\{C_{s1}, C_{s2}, \dots, C_{sq}\}$ 下计算独有属性 $\{sc_{t+1,1}, sc_{t+1,2}, \dots, sc_{t+1,q}\}$ 及其二元组结点 $\{C_{st+1}, C_{st+2}, \dots, C_{st+q}\}$ 。

如果上层结点 $\{C_{m1}, C_{m2}, \dots, C_{mp}\}$ 与其共有属性结点 $\{m_{t+1,1}, m_{t+1,2}, \dots, m_{t+1,l}\}$ 和独有属性结点 $\{C_{s1}, C_{s2}, \dots, C_{sk}\}$ 间满足关系:

$$\bigcup_{i=1}^p g(m_{t+1,i}) \cup \bigcup_{i=1}^q g(s_{t+1,i}) = g(m_j), j \in (1:p)$$

则进入下一步; 不满足时, 在最底层结点与相应共有属性的结点间绘制 l 条有向边,

$$l = \left| g(m_j) - \bigcup_{i=1}^p g(m_{t+1,i}) \cup \bigcup_{i=1}^q g(s_{t+1,i}) \right|, j \in (1:p)$$

Step 4: 当在共有属性或独有属性下不能生成新的二元结点时停止, 否则调至 Step 3。

Step 5: 算法结束。

3. 属性偏序结构图中的关联结点和关联路径

本节将在属性偏序结构图中给出关联规则的关联结点, 关联路径的相关定义及关联规则支持度和置信度的计算。

定义 6 在属性偏序结构图 $\mathfrak{R}(U, A, I)$ 中, 若结点 (A', U') 蕴涵关联规则 $X \Rightarrow Y$, 即 $X \subset A', Y \subset A'$, 则称结点 (A', U') 为关联规则 $X \Rightarrow Y$ 的关联结点。由结点 (A', U') 向底层结点 (A, \emptyset) 建立的边称作关联规则 $X \Rightarrow Y$ 的关联路径。

定义 7 对于关联规则 $X \Rightarrow Y$, 在属性偏序结构图 $\mathfrak{R}(U, A, I)$ 中的所有关联结点的集合记为 $C_{(X \Rightarrow Y)} = \{(A_1, U_1), (A_2, U_2), \dots, (A_k, U_k) | k \in N^+\}$, 定义关联规则 $X \Rightarrow Y$ 的支持度为

$$\text{support}(X \Rightarrow Y) = \frac{|U_1 \cup U_2 \cup \dots \cup U_k|}{|U|}$$

其中, $|\bullet|$ 表示基数。

定义 8 在属性偏序结构图 $\mathfrak{R}(U, A, I)$ 中, 所有包含属性集(项目集) B 的结点记为 $C_{(B)} = \{(B_1, U_{B1}), (B_2, U_{B2}), \dots, (B_n, U_{Bn}) | n \in N^+\}$, 定义 B 的支持度为

$$\text{support}(B) = \frac{|U_{B1} \cup U_{B2} \cup \dots \cup U_{Bn}|}{|U|}$$

如果 $\text{support}(B)$ 大于支持度阈值 \min_sup , 则称 B 为频繁项目集。如果 B 包含 k 个属性, 则称 B 为频繁 k -项集。

定义 9 若在属性偏序结构图 $\mathfrak{R}(U, A, I)$ 中存在关联规则 $X \Rightarrow Y$, 则其置信度为

$$\text{confidence}(X \Rightarrow Y) = \frac{\text{support}(X \cup Y)}{\text{support}(X)}$$

给定最小置信度 \min_conf , 如果规则 $X \Rightarrow Y$ 同时满足最小支持度 \min_sup 和最小置信度 \min_conf , 则称该规则为强关联规则。

定义 10 对于属性集(项目集) $W \subset X \subseteq A$, 如果 $\text{support}(W) = \text{support}(X) \geq \min_sup$, 并且对于任何的 $Z \supset X$, 均有 $\text{support}(Z) < \text{support}(X)$, 则称 X 为 W 在相同支持度下的最大频繁项目集。

定理 1 对于规则 $X \Rightarrow Y$, $W \Rightarrow Y$, 满足 $W \subset X$ 且 $\text{support}(W) = \text{support}(X)$, 则 $W \Rightarrow Y$ 是 $X \Rightarrow Y$ 的冗余规则。

证明: 因为 $\text{support}(W \Rightarrow Y) = \text{support}(W \cup Y)$, $\text{support}(X \Rightarrow Y) = \text{support}(X \cup Y)$ 且 $W \subset X$, 所以 $\text{support}(W \cup Y) \geq \text{support}(X \cup Y)$, 即 $\text{support}(W \Rightarrow Y) \geq \text{support}(X \Rightarrow Y)$ 。因此

$\frac{\text{support}(W \Rightarrow Y)}{\text{support}(W)} \geq \frac{\text{support}(X \Rightarrow Y)}{\text{support}(X)}$, 即 $\text{confidence}(W \Rightarrow Y) \geq \text{confidence}(X \Rightarrow Y)$ 。若 $X \Rightarrow Y$ 满足支持度阈值 \min_sup 和最小置信度 \min_conf , $W \Rightarrow Y$ 也必定满足条件。

定理 2 对于规则 $Y \Rightarrow X$, $Y \Rightarrow W$, 满足 $W \subset X$ 且 $\text{support}(W) = \text{support}(X)$, 则 $Y \Rightarrow W$ 是 $Y \Rightarrow X$

的冗余规则。

证明: 因为 $support(Y \Rightarrow W) = support(Y \cup W)$, $support(Y \Rightarrow X) = support(Y \cup X)$ 且 $W \subset X$, 所以 $support(Y \cup W) \geq support(Y \cup X)$, 即 $support(Y \Rightarrow W) \geq support(Y \Rightarrow X)$ 。因此 $\frac{support(Y \Rightarrow X)}{support(Y)} \geq \frac{support(Y \Rightarrow W)}{support(Y)}$, 即 $confidence(Y \Rightarrow W) \geq confidence(Y \Rightarrow X)$ 。若 $Y \Rightarrow X$ 满足支持度阈值 min_sup 和最小置信度 min_conf , $Y \Rightarrow W$ 也必定满足条件。

由定理 1 和定理 2 可知, 从数据集中挖掘相同支持度下的最大频繁项目集进行关联规则提取即可获得无冗余关联规则。

4. 基于属性偏序结构图提取关联规则的原理及方法

在 Apriori 算法中, 频繁项的提取是关联规则提取的关键步骤, Apriori 算法的总体性能主要由这一步决定。而由数据集产生的频繁项集数目往往会非常大, 导致挖掘过程比较繁琐, 生成大量冗余规则。本节提出基于属性偏序结构图的关联规则提取方法, 从属性偏序结构图中寻找相同支持度下的最大频繁项目集, 减少频繁项目集数量, 提高关联规则挖掘效率, 并获得无冗余的关联规则。下面从构图原理和数学角度说明从属性偏序结构图中提取频繁项的合理性和科学性。

根据属性偏序结构图的构图步骤 step 2, 计算最大共有属性 M_0 下所有共有属性集合 $\{m_{11}, m_{12}, \dots, m_{1p}\}$ 和独有属性集合 $\{s_{11}, s_{12}, \dots, s_{1q}\}$ 及其对应二元组 $\{C_{m1}, C_{m2}, \dots, C_{mp}\}$ 和 $\{C_{s1}, C_{s2}, \dots, C_{sq}\}$, 并与最大共有属性的结点 C_0 间绘制有向边, 说明属性偏序结构图是以属性覆盖对象的程度为偏序进行构图, 所以相同支持度下的最大频繁项目集必定含于属性偏序结构图的结点中。如果该层有序对结点与最大共有属性的结点间满足关系 $\bigcup_{i=1}^p g(m_{1i}) \cup \bigcup_{i=1}^q g(s_{1i}) = U$, 则进入下一步。这保证了覆盖对象的全面性。也就是说形式背景相同支持度下的最大频繁项目集必定包含属性集 $\{m_{11}, m_{12}, \dots, m_{1p}\}$ 或 $\{s_{11}, s_{12}, \dots, s_{1q}\}$ 中的元素。属性偏序结构图体现的是自下而上的属性聚合模式, 所以可以从上到下扫描属性偏序结构图提取频繁项, 即可得到相同支持度下的最大频繁项目集, 进而提取无冗余的关联规则。

本文提出的基于属性偏序结构图的关联规则提取按照以下步骤进行。

Step 1: 设置最小支持度 min_sup , 对形式背景进行第一遍扫描, 生成频繁 1-项集;

Step 2: 由形式背景生成对应的属性偏序结构图, 记最大共有属性 M_0 下所有共有属性集合和独有属性集合分别为 $\{m_{11}, m_{12}, \dots, m_{1p}\}$ 和 $\{s_{11}, s_{12}, \dots, s_{1q}\}$;

Step 3: 在属性偏序结构图中由上到下依次寻找包含 $\{m_{11}, m_{12}, \dots, m_{1p}\}$ 或 $\{s_{11}, s_{12}, \dots, s_{1q}\}$ 的频繁项集并提取对象支持集;

Step 4: 设置最小支持度 min_conf , 依据挖掘的频繁项目集提取强关联规则并计算相应的置信度。

5. 例子说明

例 1. 表 1 给出了一个形式背景 (U, A, I) , 对象集(事物集) $U = \{1, 2, 3, 4, 5, 6\}$, 属性集 $A = \{a, b, c, d, e, f, g\}$ 。它们之间的二元关系 I 如表 1 所示。表中的 1 表示对象具有某属性, 0 表示对象不具有某属性。利用本文所提方法进行关联规则提取步骤如下。

Step 1: 设最小支持度 $min_sup = 2/7$, 对形式背景进行第一遍扫描, 得到频繁 1-项集, 见表 2。

Step 2: 由表 1 给出的形式背景生成对应的属性偏序结构图 \mathfrak{R} , 如图 1 所示。

Step 3: 提取属性偏序结构图中的频繁项。由图 1 可知结点 $(a, 1234)$, $(c, 6)$ 和 $(d, 5)$ 位于属性偏序结构

Table 1. A formal context (U, A, I)

表 1. 形式背景 (U, A, I)

$X \setminus A$	a	b	c	d	e
1	1	1	1	0	0
2	1	1	1	0	1
3	1	0	1	0	0
4	1	1	0	0	1
5	0	0	0	1	0
6	0	0	1	0	0
7	1	1	0	1	0

Table 2. Frequent 1-items

表 2. 频繁 1-项集

频繁项集	支持度
a	5/7
b	4/7
c	4/7
d	2/7
e	2/7

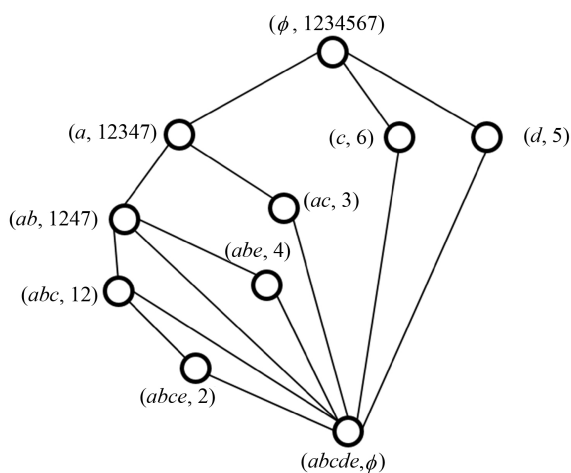


Figure 1. APOSD \mathfrak{R}

图 1. 属性偏序结构图 \mathfrak{R}

\mathfrak{R} 的第二层, 对应的属性分别为 a, c, d 。在图 1 中逐层向下搜索寻找相同支持度下的最大频繁项集, 得到相同支持度下的最大频繁项集及其支持度如图 2 所示。

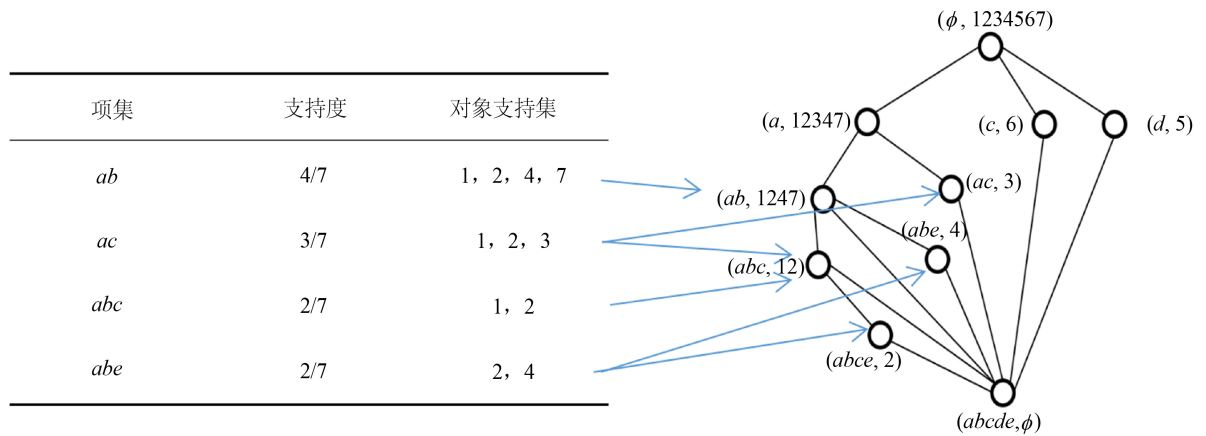


Figure 2. The frequent items are extracted from the APOSD
图 2. 从属性偏序结构图中提取的频繁项集

Step 4: 设 $min_conf = 80\%$, 结合表 1 表 2 的频繁项集进行关联规则提取, 得到关联规则, 并计算其置信度如下:

$$b \Rightarrow a, \quad confidence(b \Rightarrow a) = \frac{support(ab)}{support(b)} = 100\%;$$

$$a \Rightarrow b, \quad confidence(a \Rightarrow b) = \frac{support(ab)}{support(a)} = 80\%;$$

$$bc \Rightarrow a, \quad confidence(bc \Rightarrow a) = \frac{support(abc)}{support(bc)} = 100\%;$$

$$e \Rightarrow ab, \quad confidence(e \Rightarrow ab) = \frac{support(abe)}{support(e)} = 100\%。$$

6. 结果对比分析

关联规则挖掘方法中, 最典型的关联规则提取方法是 Apriori 算法。它是一种逐层搜索的迭代方法。扫描数据库直接得到频繁 1-项集, 在得到频繁 1-项集的基础上寻找频繁 2-项集。如此循环, 在得到 k -项集的基础上得到 $(k + 1)$ -项集, 直到无法找出高频项集, 停止迭代。根据频繁项集, 导出所有的关联规则。

在设置同样的最小支持度 min_sup 和最小置信度 min_conf 时, 利用传统的 Apriori 算法产生的频繁项集及支持度如表 3 所示。

根据表 3 中的频繁项集, 提取出的关联规则及置信度如表 4 所示。

对比可知, Apriori 算法得到频繁项目集为 11 个, 在属性偏序结构图中利用本文提出的方法得到的频繁项目集为 9 个。频繁项目集个数缩减了 2 个, 在一定程度上起到了减少提取关联规则计算时间的作用。基于属性偏序结构图的关联规则提取方法获得的关联规则为 $b \Rightarrow a, a \Rightarrow b, bc \Rightarrow a, e \Rightarrow ab$ 。与 Apriori 算法比较减少 2 个关联规则 $e \Rightarrow a, e \Rightarrow b$ 。而关联规则 $e \Rightarrow a, e \Rightarrow b$ 是所获得的关联规则 $e \Rightarrow ab$ 的冗余规则, 可根据定理 2 证明。

Apriori 算法的不足之处在于, 每次寻求频繁 k -项集都需要扫描一次数据库, 需要对数据库进行反复扫描, 占用了大量的时间和空间, 并且生成大量的冗余规则。本文提出的方法通过对属性偏序结构图的频繁模式的遍历操作, 代替了 Apriori 算法中对形式背景的多次扫描。在提取的频繁项目集的过程中,

Table 3. Frequent items and support degrees are based on Apriori algorithm
表 3. Apriori 算法产生的频繁项集及支持度

频繁项集	支持度	频繁项集	支持度
a	5/7	ac	3/7
b	4/7	ae	2/7
c	4/7	be	2/7
d	2/7	abc	2/7
e	2/7	abe	2/7
ab	4/7		

Table 4. Association rules and confidence are based on Apriori algorithm
表 4. Apriori 算法产生的关联规则及置信度

关联规则	置信度	关联规则	置信度
$b \Rightarrow a$	100%	$e \Rightarrow ab$	100%
$a \Rightarrow b$	80%	$e \Rightarrow a$	100%
$bc \Rightarrow a$	100%	$e \Rightarrow b$	100%

获得的频繁项目集是相同支持度下的最大频繁项目集, 从而在属性偏序结构图中寻找频繁项目集比 Apriori 算法挖掘频繁项目集个数有所缩减, 同时更具有代表性, 在已有的频繁项目集下生成的关联规则是无冗余的关联规则。除此之外, 本文提出的基于属性偏序结构图的关联规则提取方法, 使关联关系通过属性偏序结构图的结点的泛化与特化关系直观地体现出来, 可读性比较强。如关联规则 $b \Rightarrow a$, 对应于属性偏序结构图中的关联结点为 $(ab, 1247)$, $(abc, 12)$, $(abce, 2)$, $(abe, 4)$, 标记出对应关联结点和关联路径, 有助于用户理解所提取的关联规则, 便于对关联规则进行深入分析。

7. 总结

本文首先证明了提取相同支持度下的最大频繁项目集即可得到无冗余关联规则。在分析属性偏序结构图的构图原理的基础上, 从数学角度说明在属性偏序结构图中提取相同支持度下的最大频繁项目集的合理性和科学性, 进而提出了基于属性偏序结构图的关联规则提取方法, 获得无冗余的关联规则, 实现频繁项分层的关联规则可视化展示。结合对象和属性两个角度从全局的知识背景中去分析关联关系, 提高了对潜在知识和非关联关系的利用和发掘程度。将关联规则转换成属性偏序结构图中的知识表示形式, 具有较强的可读性, 有助于用户对关联规则进行深入分析和研究。

参考文献

- [1] Wille, R. (1982) Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts. *Orderd Sets D Reidel*, **83**, 314-339. https://doi.org/10.1007/978-3-642-01815-2_23
- [2] Agrawal, R., Imieliński, T. and Swami, A. (1993) Mining Association Rules between Sets of Items in Large Databases. *ACM SIGMOD Record*, **22**, 207-216. <https://doi.org/10.1145/170036.170072>
- [3] Feng, H., Liao, R.T., Liu, F., et al. (2018) Optimization Algorithm Improvement of Association Rule Mining Based on Particle Swarm Optimization. *International Conference on Measuring Technology & Mechatronics Automation*, IEEE Computer Society, 524-529. <https://doi.org/10.1109/ICMTMA.2018.00132>
- [4] Wang, D.X., Xie, Q., Huang, D.M., et al. (2012) Analysis of Association Rule Mining on Quantitative Concept Lattice. *International Conference on Artificial Intelligence & Computational Intelligence*, Springer, Berlin, Heidelberg,

142-149. https://doi.org/10.1007/978-3-642-33478-8_19

- [5] 胡可云, 陆玉昌, 石纯一. 基于概念格的分类和关联规则的集成挖掘方法[J]. 软件学报, 2000(11): 1478-1484.
- [6] 王德兴, 胡学钢, 刘晓平, 等. 基于概念格和 Apriori 的关联规则挖掘算法分析[J]. 合肥工业大学学报(自然科学版), 2006(6): 699-702.
- [7] 杨葛英, 沈夏炯, 史先进, 等. 以概念格为背景的关联规则可视化[J]. 计算机工程与应用, 2021, 57(1): 84-91.
- [8] 洪文学, 李少雄, 张涛, 栾景民, 刘文远. 大数据偏序结构生成原理[J]. 燕山大学学报, 2014, 38(5): 388-393+402.
- [9] 刘倩, 赵岩松, 洪文学, 等. 基于属性偏序结构图对《妇人大全良方》中治疗胎停育相关方药分析[J]. 现代中西医结合杂志, 2017, 26(8): 799-802, 815.
- [10] 顾广华, 曹宇尧, 李刚, 等. 基于语义标签生成和偏序结构的图像层级分类[J]. 软件学报, 2020, 31(2): 289-301.
- [11] 徐笋晶, 李赛美, 洪文学, 等. 基于数学属性偏序表示原理的李赛美教授治疗 2 型糖尿病处方用药分析[J]. 中国实验方剂学杂志, 2014, 20(2): 207-211.
- [12] 徐伟华, 等. 形式概念分析理论与应用[M]. 北京: 科学出版社, 2016.