

# 带对象结构信息形式背景的概念知识发现与演化

闫梦宇<sup>1,2</sup>, 李金海<sup>1,2</sup>, 刘文奇<sup>1,2</sup>, 张文修<sup>3</sup>

(1. 昆明理工大学数据科学研究中心, 云南昆明 650500;  
2. 昆明理工大学理学院, 云南昆明 650500; 3. 西安交通大学数学与统计学院, 陕西西安 710049)

**摘要:** 形式概念分析是数据处理和知识发现的有力工具, 然而仅利用形式概念分析中的传统方法已不能满足大数据环境的需求, 这是因为大量数据不仅包括对象-属性信息, 还有对象与对象之间的结构信息, 因此扩展形式概念分析进行数据处理与知识发现是一个重要的研究课题. 基于此, 本文定义了带对象结构信息形式背景, 并在形式概念的基础上引入连通性提出全局结构信息形式概念和局部结构信息形式概念(统称为结构信息形式概念), 进而讨论基于结构信息形式概念的知识发现问题; 由于现实中数据的动态变化及信息的传播与扩散是不可避免的, 又进一步探讨了当带对象结构信息形式背景对象结构矩阵或连接矩阵发生变化时结构信息形式概念的演化规律, 根据不同情况分别给出了对应的算法并对它们的时间复杂度进行分析; 此外, 通过实验表明了结构信息形式概念动态更新方法的可行性与有效性.

**关键词:** 形式概念分析; 结构信息形式概念; 知识发现; 概念演化; 动态更新

**基金项目:** 国家自然科学基金(No.11971211, No.12171388)

**中图分类号:** TP182

**文献标识码:** A

**文章编号:** 0372-2112(2023)01-0011-07

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20211559

## Conceptual Knowledge Discovery and Evolution in Formal Context with Object Structure Information

YAN Meng-yu<sup>1,2</sup>, LI Jin-hai<sup>1,2</sup>, LIU Wen-qi<sup>1,2</sup>, ZHANG Wen-xiu<sup>3</sup>

(1. *Data Science Research Center, Kunming University of Science and Technology, Kunming, Yunnan 650500, China;*  
2. *Faculty of Science, Kunming University of Science and Technology, Kunming, Yunnan 650500, China;*  
3. *School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China*)

**Abstract:** Formal concept analysis is a powerful tool for data processing and knowledge discovery. However, only using the traditional methods in formal concept analysis cannot meet the needs of big data environment since most data not only contain object-attribute information but also the structure information between objects. As a result, it is an important research topic to extend formal concept analysis for data processing and knowledge discovery. Based on the above analysis, this paper puts forward the notion of a formal context with object structure information. And then, the global structural information and local structural information formal concepts are proposed (collectively called structural information formal concepts) by combining with connectivity. Meanwhile, the problem of knowledge discovery on the basis of structural information formal concepts is discussed. Considering the fact that the dynamic changes of data and the dissemination or diffusion of information in our daily life are inevitable, we further explore the evolution rules of structural information formal concepts when a formal context with object structure information is updated gradually. Algorithms are developed for updating different structural information formal concepts when object structure matrix or connection matrix changes, and their time complexity are also analyzed. In addition, we conduct some experiments to show the feasibility and effectiveness of the proposed structural information formal concept updating methods.

**Key words:** formal concept analysis; structural information formal concept; knowledge discovery; concept evolution; dynamic updating

**Foundation Item(s):** National Natural Science Foundation of China (No.11971211, No.12171388)

## 1 引言

知识发现是从大量数据中抽取有价值的知识的过程<sup>[1,2]</sup>. 而概念是知识表示的一种有效手段,从数据库中发现概念知识是指将数据库中蕴含的知识形式化为有用的概念,这一过程称为概念知识发现,对处理和解释数据十分有益.

形式概念分析(Formal Concept Analysis, FCA)是一种实用的概念知识发现工具,它是 Wille<sup>[3]</sup>教授于1982年提出的,研究对象是一类特定的对象-属性关系数据库(又称为形式背景). 通过对象集与属性集之间的特定关联,建立由对象集和属性集序对形成的概念知识. 近年来, FCA 已被视为是“概念数据分析与知识表示”的重要数学工具,在知识发现<sup>[4]</sup>、机器学习<sup>[5,6]</sup>、数据挖掘<sup>[7]</sup>、软件工程<sup>[8]</sup>、特征选择<sup>[9,10]</sup>、概念认知<sup>[11,12]</sup>、异构数据<sup>[13]</sup>、属性拓扑<sup>[14]</sup>、决策蕴涵<sup>[15]</sup>等领域得到了广泛应用.

FCA 主要处理对象集与属性集之间的关系数据库,忽略了对对象与对象之间的信息. 早期在 FCA 领域考虑对象-对象关系的相关成果是通过对象-对象形式背景和对象-属性形式背景讨论对象和属性之间的联系<sup>[16]</sup>,但并没有真正将对象之间的结构信息考虑在内. 基于上述讨论,本文引入连通性定义了带对象结构信息形式背景,需要指出的是,一个形式概念可视为一类,在形式概念的基础上考虑对象之间的连通性只能判断给定带对象结构信息形式背景中具有相同属性的对象是否属于同一类,这相当于全局连通性;然而,给定带对象结构信息形式背景中具有相同属性的部分连通对象构成同一类的现象也是经常存在的,若只考虑全局连通性则局部包含结构信息的概念知识将无法获得. 因此,本文将对象的结构信息进一步又区分为全局信息与局部信息,提出2种结构信息形式概念知识发现模型.

不仅如此,数据也会随时间不断动态更新<sup>[17]</sup>. 在 FCA 领域,数据的动态更新主要指对象个数、属性个数以及属性值的变化<sup>[18]</sup>. 对于这些不断动态更新的数据,如何快速准确地找到特定的知识受到了众多学者的关注<sup>[19,20]</sup>. 受此启发,本文将进一步提出动态概念知识发现方法以探讨概念演化问题.

## 2 带对象结构信息形式背景的概念发现

对于形式背景  $(G, M, C)$ ,  $X \subseteq G, B \subseteq M$ , 定义算子

$$X^* = \{m \in M \mid \forall g \in X, (g, m) \in C\}$$

$$B^* = \{g \in G \mid \forall m \in B, (g, m) \in C\}$$

若  $X^* = B$  且  $B^* = X$ , 则称  $(X, B)$  为形式概念, 其中  $X$  为形式概念的外延,  $B$  为形式概念的内涵.

**定义 1** 四元组  $(G, M, A, C)$  称为带对象结构信息形式背景, 其中  $G = \{g_1, g_2, \dots, g_n\}$  是非空有限对象集,  $M = \{m_1, m_2, \dots, m_p\}$  是非空有限属性集,  $A$  是  $G$  上的结构矩阵, 即对任意的  $g_i, g_j \in G, a_{ij} = 1$  表示对象  $g_i, g_j$  是相关联的,  $a_{ij} = 0$  表示对象  $g_i, g_j$  不是相关联的;  $C$  是  $G \times M$  上的连接矩阵, 即任意的  $g_i \in G$  和  $m_k \in M, (g_i, m_k) = 1$  表示对象  $g_i$  拥有属性  $m_k, (g_i, m_k) = 0$  表示对象  $g_i$  不拥有属性  $m_k$ .

一个图称为是连通的, 当且仅当从任一节点出发, 经过一个或多个中间节点, 能够达到其余任意节点.

**定义 2** 设  $(G, M, A, C)$  为带对象结构信息形式背景, 若  $X \subseteq G$  包含的所有节点由它们之间的结构信息诱导出的子图是连通的, 则称  $X$  是连通的.

**定义 3** 设  $(G, M, A, C)$  为带对象结构信息形式背景,  $X \subseteq G, B \subseteq M (X \neq \emptyset \text{ 且 } B \neq \emptyset)$ . 若  $X^* = B, B^* = X$  且  $X$  连通, 则称  $(X, B)$  为全局结构信息形式概念. 记所有全局结构信息形式概念构成的集合为  $N(G, M, A, C)$ .

**定义 4** 设  $(G, M, A, C)$  为带对象结构信息形式背景,  $X \subseteq G, B \subseteq M (X \neq \emptyset \text{ 且 } B \neq \emptyset)$ . 若  $X^* = B, X$  连通且不存在  $g \in B^* - X$  使得  $X \cup \{g\}$  连通, 则称序对  $(X, B)$  为局部结构信息形式概念. 记所有局部结构信息形式概念构成的集合为  $N_L(G, M, A, C)$ .

为了方便, 将全局和局部结构信息形式概念统称为结构信息形式概念.

算法 1 和算法 2 分别给出了从给定线索中发现相应的全局和局部结构信息形式概念的具体过程.

---

### 算法 1 全局结构信息形式概念发现

---

输入: 带对象结构信息形式背景  $(G, M, A, C)$  和线索集  $X$

输出: 全局结构信息形式概念  $(X, B)$

1. IF  $X^{**}$  连通且  $X^* \neq \emptyset$  THEN
  2. 令  $X = X^{**}, B = X^*$ , 返回  $(X, B)$ ;
  3. ELSE
  4. 返回  $(X, B) = \emptyset$ ;
  5. END IF
- 

### 算法 2 局部结构信息形式概念发现

---

输入: 带对象结构信息形式背景  $(G, M, A, C)$  和线索集  $X$

输出: 局部结构信息形式概念  $(X, B)$

1. FOR  $X^{**}$  的每个最大连通子集  $X_1$
  2. 令  $X = X_1, B = X_1^*$ ;
  3. IF  $B \neq \emptyset$  THEN
  4. 返回  $(X, B)$ ;
  5. ELSE
  6. 返回  $(X, B) = \emptyset$ ;
  7. END IF
  8. END FOR
-

分析可得,算法1的时间复杂度为 $O(|X||G|(|A|+|G|))$ ,算法2的时间复杂度为 $O(|X||A||G|^2)$ .

### 3 结构信息形式概念的演化

本节讨论对象结构信息或对象-属性信息发生改变时,结构信息形式概念的演化情况.

#### 3.1 对象结构矩阵改变时现有结构信息形式概念的演化

此时只有对象结构矩阵 $A$ 发生变化,设其更新为 $A'$ .为了方便,记 $(G, M, A', C)$ 的所有全局结构信息形式概念为 $N(G, M, A', C)$ .根据定义3可得如下性质.

**性质1** 设 $(G, M, A, C)$ 为带对象结构信息形式背景, $a_{ij}=a_{ji}$ 从1变为0,那么 $(X, B) \in N(G, M, A, C)$ 满足:

- (1) 若 $X$ 依旧连通,则 $(X, B) \in N(G, M, A', C)$ ;
- (2) 若 $X$ 不连通,对于 $X$ 的任一连通子集 $X_1$ ,得到新的序对 $(X_1, X_1^*)$ .如果 $X_1^* \supset B$ 且 $X_1^{**}=X_1$ ,那么 $(X_1, X_1^*) \in N(G, M, A', C)$ ;否则 $(X_1, X_1^*) \notin N(G, M, A', C)$ .

为了验证该性质的可行性,将其操作过程记为算法3,其时间复杂度为 $O(|G|+|X|^2)$ .类似可得下列性质.

#### 算法3 减边时全局结构信息形式概念的演化

输入:带对象结构信息形式背景 $(G, M, A', C)$ 和 $(X, B) \in N(G, M, A, C)$

输出:演化后新全局结构信息形式概念 $(X_{\text{new}}, B_{\text{new}})$

1. IF  $X$  连通 THEN
2. 令 $X_{\text{new}}=X, B_{\text{new}}=B$ ,返回 $(X_{\text{new}}, B_{\text{new}})$ ;
3. ELSE
4. 计算 $X$ 的连通子集 $X_1$ ;
5. END IF
6. IF  $X_1^*=B$ 或 $\{X_1^* \supset B$ 且 $X_1^{**} \neq X_1\}$  THEN
7. 返回 $(X_{\text{new}}, B_{\text{new}})=\emptyset$ ;
8. ELSE
9. 令 $X_{\text{new}}=X_1, B_{\text{new}}=X_1^*$ ,返回 $(X_{\text{new}}, B_{\text{new}})$ ;
10. END IF

**性质2** 设 $(G, M, A, C)$ 为带对象结构信息形式背景, $(X, B) \in N(G, M, A, C)$ ,若 $a_{ij}=a_{ji}$ 从0变为1,则 $(X, B) \in N(G, M, A', C)$ .

下面讨论对象结构矩阵改变后,局部结构信息形式概念的演化情况.记 $(G, M, A', C)$ 中的所有局部结构信息形式概念为 $N_L(G, M, A', C)$ .根据定义4可得下列性质.

**性质3** 设 $(G, M, A, C)$ 为带对象结构信息形式背景, $a_{ij}=a_{ji}$ 从1变为0,那么 $(X, B) \in N_L(G, M, A, C)$ 满足:

- (1) 若 $X$ 在 $(G, M, A', C)$ 中连通,则 $(X, B) \in N_L(G, M, A', C)$ ;
- (2) 若 $X$ 在 $(G, M, A', C)$ 中不连通,则 $(X, B) \notin N_L(G, M, A', C)$ .对 $X$ 的最大连通子集 $X_0$ ,有

$(X_0, X_0^*) \in N_L(G, M, A', C)$ .

同理,为了验证该性质的可行性,需要给出它的实现算法,但考虑到其与全局结构信息形式概念的演化类似,故只记其为算法3',并省略具体细节.

**性质4** 设 $(G, M, A, C)$ 为带对象结构信息形式背景, $a_{ij}=a_{ji}$ 从0变为1,则 $(X, B) \in N_L(G, M, A, C)$ 满足:

- (1) 如果存在 $g \in B^* - X$ 在 $(G, M, A', C)$ 中与 $X$ 连通,那么 $(X, B) \notin N_L(G, M, A', C)$ ;
- (2) 如果不存在 $g \in B^* - X$ 在 $(G, M, A', C)$ 中与 $X$ 连通,那么 $(X, B) \in N_L(G, M, A', C)$ .

#### 3.2 对象结构矩阵改变时形式概念或弱概念演化成结构信息形式概念

对于非结构信息形式概念 $(X, B)$ ,当带对象结构信息形式背景减少边时,它对 $(X, B)$ 能否提升为结构信息形式概念没有影响,故文中只讨论增加边的情况.下列性质可由定义3和定义4得到.

**性质5** 设 $(G, M, A, C)$ 为带对象结构信息形式背景, $a_{ij}=a_{ji}$ 从0变为1, $(X, B) \in L(G, M, C) - N(G, M, A, C)$ 满足:

- (1) 若 $X$ 在 $(G, M, A', C)$ 中连通,则 $(X, B) \in N(G, M, A', C)$ ;
- (2) 若 $X$ 在 $(G, M, A', C)$ 中不连通,则 $(X, B) \notin N(G, M, A', C)$ .

**性质6** 设 $(G, M, A, C)$ 为带对象结构信息形式背景, $a_{ij}=a_{ji}$ 从0变为1, $(X, B) \in L(G, M, C) - N_L(G, M, A, C)$ ,则下列结论成立:

- (1) 如果 $X$ 在 $(G, M, A', C)$ 中连通,那么 $(X, B) \in N_L(G, M, A', C)$ ;
- (2) 如果 $X$ 在 $(G, M, A', C)$ 中不连通,那么 $(X, B) \notin N_L(G, M, A', C)$ .

需要指出的是,当带对象结构信息形式背景增加边时,虽然序对 $[X, B]$ (满足 $X^*=B$ 且 $B^* \neq X$ )不是全局结构信息形式概念,但很可能成为局部结构信息形式概念.因此,接下来讨论序对 $[X, B]$ 演化成局部结构信息形式概念的情况.

**定义5** 设 $(G, M, A, C)$ 为带对象结构信息形式背景,若 $[X, B]$ 满足 $X^*=B$ 且 $B^* \neq X$ ,则称 $[X, B]$ 为弱概念.

**性质7** 设弱概念 $[X, B] \notin N(G, M, A, C)$ ,如果 $X$ 在 $(G, M, A', C)$ 中连通且不存在 $g \in B^* - X$ 与 $X$ 连通,则 $[X, B] \in N_L(G, M, A', C)$ ;否则 $[X, B] \notin N_L(G, M, A', C)$ .

为了验证当对象结构矩阵改变时形式概念演化成全局结构信息形式概念的计算效率,将性质5的具体实施步骤记为算法4,它的时间复杂度为 $O(|X|^2)$ .对于当对象结构矩阵改变时形式概念或弱概念演化成局部结

构信息形式概念的计算效率,也可给出类似的算法,但为了节省篇幅,记其为算法4',并省略细节.

#### 算法4 加边时形式概念演化为全局结构信息形式概念

输入:带对象结构信息形式背景 $(G, M, A', C)$ 和 $(X, B) \in L(G, M$

$C) - N(G, M, A, C)$

输出:演化后新全局结构信息形式概念 $(X_{\text{new}}, B_{\text{new}})$

1. IF  $X$  连通 THEN
2. 令  $X_{\text{new}} = X, B_{\text{new}} = B$ , 返回  $(X_{\text{new}}, B_{\text{new}})$ ;
3. ELSE
4. 返回  $(X_{\text{new}}, B_{\text{new}}) = \emptyset$ ;
5. END IF

### 3.3 连接矩阵改变时结构信息形式概念的演化

为了方便,记 $C'$ 为 $C$ 更新后的连接矩阵,新的带对象结构信息形式背景为 $(G, M, A, C')$ .

**定义6** 设 $(G, M, A, C)$ 为带对象结构信息形式背景,连接矩阵 $C$ 更新后得到的形式背景为 $(G, M, A, C')$ ,其诱导算子为

$$X^{*C'} = \{m \in M | \forall g \in X, (g, m) \in C'\}$$

$$B^{*C'} = \{g \in G | \forall m \in B, (g, m) \in C'\}$$

由于连接矩阵中属性值改变时结构信息形式概念演化情况与经典形式概念动态更新类似,故相关性不再赘述,只给出其演化算法.算法5和6分别是属性值由0变为1以及从1变为0时全局结构信息形式概念的演化算法,它们的时间复杂度分别为 $O(|X|(|A| + |G| + |X|^2))$ 和 $O(|X|(|A| + |X|^2))$ .另外,记与之对应的局部结构信息形式概念的演化为算法5'和算法6',并省略细节.

#### 算法5 $(g_i, m_j) = 0$ 变为1时全局结构信息形式概念的演化

输入:带对象结构信息形式背景 $(G, M, A, C'), (X, B) \in N(G, M,$

$A, C)$ 和 $g_i, m_j$

输出:演化后新全局结构信息形式概念 $(X_{\text{new}}, B_{\text{new}})$

1. IF  $g_i \in X$  且  $m_j \notin B$  THEN
2. 令  $X_{\text{new}} = X, B_{\text{new}} = X^{*C'}$ , 返回  $(X_{\text{new}}, B_{\text{new}})$ ;
3. END IF
4. IF  $\{g_i \notin X$  且  $m_j \in B\}$  或  $\{g_i \notin X$  且  $m_j \notin B\}$  THEN
5. IF  $B \not\subseteq g_i^{*C'}$  THEN
6. 令  $X_{\text{new}} = X, B_{\text{new}} = B$ , 返回  $(X_{\text{new}}, B_{\text{new}})$ ;
7. END IF
8. IF  $B \subseteq g_i^{*C'}$  且  $X \cup \{g_i\}$  连通 THEN
9. 令  $X_{\text{new}} = X \cup \{g_i\}, B_{\text{new}} = B$ , 返回  $(X_{\text{new}}, B_{\text{new}})$ ;
10. ELSE  $\{B \subseteq g_i^{*C'}$  且  $X \cup \{g_i\}$  不连通
11. 返回  $(X_{\text{new}}, B_{\text{new}}) = \emptyset$ ;
12. END IF
13. END IF

#### 算法6 $(g_i, m_j) = 1$ 变为0时全局结构信息形式概念的演化

输入:带对象结构信息形式背景 $(G, M, A, C'), (X, B) \in N(G, M,$

$A, C)$ 和 $g_i, m_j$

输出:演化后新全局结构信息形式概念 $(X_{\text{new}}, B_{\text{new}})$

1. IF  $g_i \in X$  且  $m_j \in B$  THEN
2. 令  $X_{\text{new}} = (B - \{m_j\})^{*C'}, B_{\text{new}} = B - \{m_j\}$ ;
3. IF  $X_{\text{new}}$  连通 THEN
4. 返回  $(X_{\text{new}}, B_{\text{new}})$ ;
5. ELSE
6. 返回  $(X_{\text{new}}, B_{\text{new}}) = \emptyset$ ;
7. END IF
8. ELSE
9. 令  $X_{\text{new}} = X, B_{\text{new}} = B$ , 返回  $(X_{\text{new}}, B_{\text{new}})$ ;
10. END IF

## 4 实验

本节通过实验评估结构信息形式概念演化的有效性.一是当带对象结构信息形式背景动态更新时结构信息形式概念的演化情况;二是当对象结构信息的类型改变时结构信息形式概念的演化情况.具体为动态更新算法(算法3,4,5,6及3',4',5',6')与传统计算方法(定义3和定义4)的对比.

实验选取DBLP真实数据集中提取到的12个会议或期刊论文数据集,其中对象表示作者,属性表示作者对应论文中的关键词,对象结构信息表示作者之间存在论文合作关系.具体信息如表1所示,其中 $|G|$ 表示对象个数, $|M|$ 表示属性个数.

图1给出了当带对象结构信息形式背景动态更新(即增加边、减少边和属性值改变)时,数据集1~12“结构信息形式概念和形式概念在新形式背景中是否为全局和局部结构信息形式概念”的耗时对比情况,其中(a)~(h)分别是减边、加边时全局和局部结构信息形式概念的计算比较,属性值由0变为1、由1变为0时全局和局部结构信息形式概念的计算比较.由于数据集规模从小到大排列,从图1可以得出:随着数据集规模的增大,利用传统计算方法的时间增加幅度越来越大,但利用动态更新方法的时间增加幅度趋于平缓,且计算效率的优势较大.这表明动态更新方法比传统计算方法更高效,且数据集规模越大,优势愈明显.

图2是将数据集1~12的数据规模增加情况与文中动态更新方法的耗时增加情况分别提取出来,以双Y轴点线图的方式予以呈现,其中相同颜色的点线对应相同颜色的Y坐标轴,(a)~(h)分别是针对减边、加边情形的全局和局部结构信息形式概念,属性值由0变为1、由1变为0时全局和局部结构信息形式概念的计算比较结果.从图2中可以看出,随着数据规模的增加幅度越来越大,利用本文动态更新方法进行概念更新所需时间

表 1 带对象结构信息形式背景真实数据集及其参数

	数据集	对象结构矩阵规模	连接矩阵规模	带对象结构信息形式背景规模	$ G $	$ M $
1	SimVis	256×256	256×205	461	256	205
2	SAGT	161×161	161×725	886	161	725
3	J Med Syst	452×452	452×792	1 244	452	792
4	MIRAGE	360×360	360×1 531	1 891	360	1 531
5	USENIX ATC	479×479	479×1 445	1 924	479	1 445
6	ICHIT	521×521	521×2 216	2 737	521	2 216
7	IEEE RTAS	735×735	735×2 421	3 156	735	2 421
8	Seville	816×816	816×3 112	3 928	816	3 112
9	ICVS	606×606	606×3 450	4 056	606	3 450
10	ACM SIGCSE Bulletin	1 042×1 042	1 042×3 814	4 856	1 042	3 814
11	ACM Trans.Math.Softw	1 870×1 870	1 870×9 751	11 621	1 870	9 751
12	IEEE Softw	4 828×4 828	4 828×16 383	21 211	4 828	16 383

的增加幅度趋于平缓,即本文动态更新方法对大规模数据集依旧适用。

最后,将数据集 1~12 的对象结构信息的类型由“作者之间存在论文合作关系”更改为“作者是否同属一个单位”,以分析本文动态更新方法受其影响的程度。图 3 是数据集 1~12 对象结构信息类型改变前后的概念演化耗时对比图,其中(a)~(h)分别是减边、加边情形的全局

和局部结构信息形式概念,属性值由 0 变为 1、由 1 变为 0 时全局和局部结构信息形式概念的演化耗时情况。从图 3 可以看出,当对象结构信息类型发生改变时,利用定义计算更新后的结构信息形式概念在最后一个数据集出现了较大差异,但是利用文中动态更新方法进行计算其差异性较小。因此,对于数据集对象结构信息类型改变的问题,动态更新方法较传统方法有较大的优势。

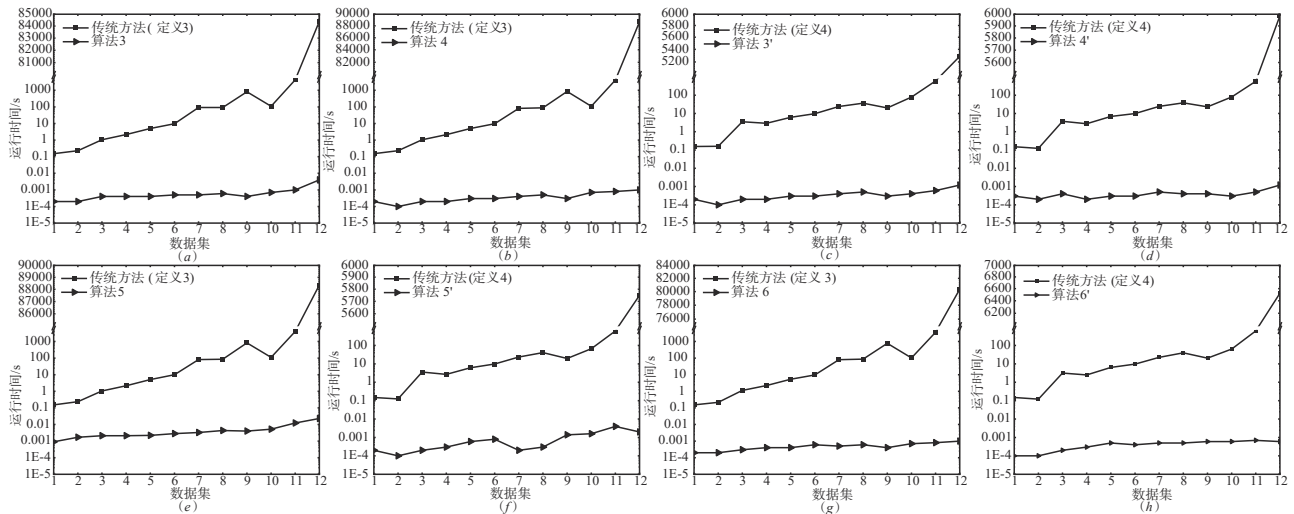


图 1 数据集 1~12 结构信息形式概念计算方法耗时对比

### 5 结束语

本文定义了带对象结构信息形式背景,基于连通性的不同理解给出了全局结构信息形式概念和局部结

构信息形式概念,进一步讨论了带对象结构信息形式背景动态更新时结构信息形式概念的变化规律,并通过实验表明对象结构信息形式概念动态更新方法的有

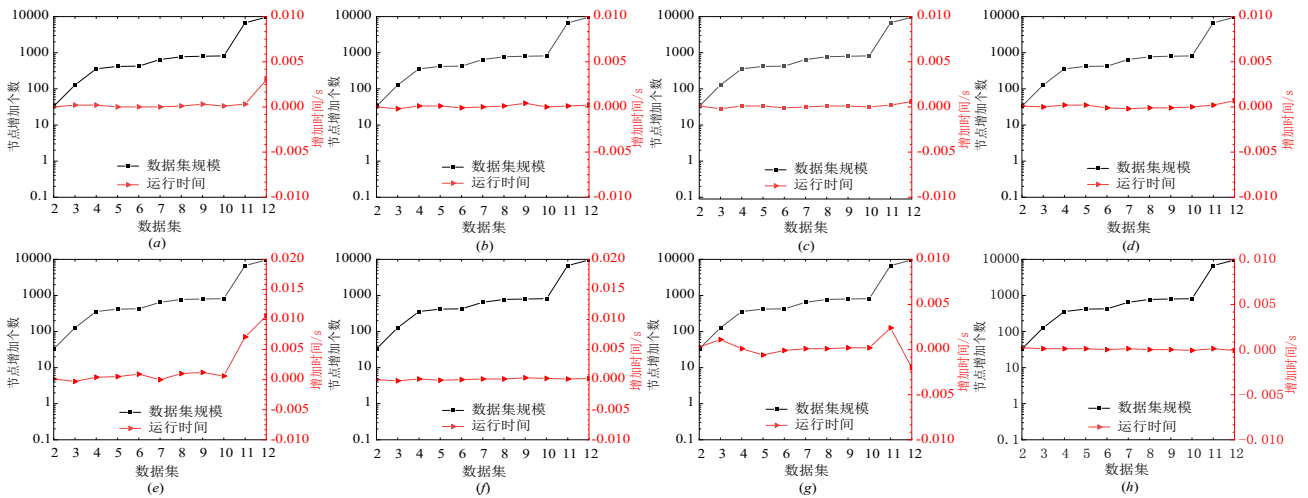


图2 数据集1~12数据规模增加情况与网络概念动态更新方法的耗时增加情况对比

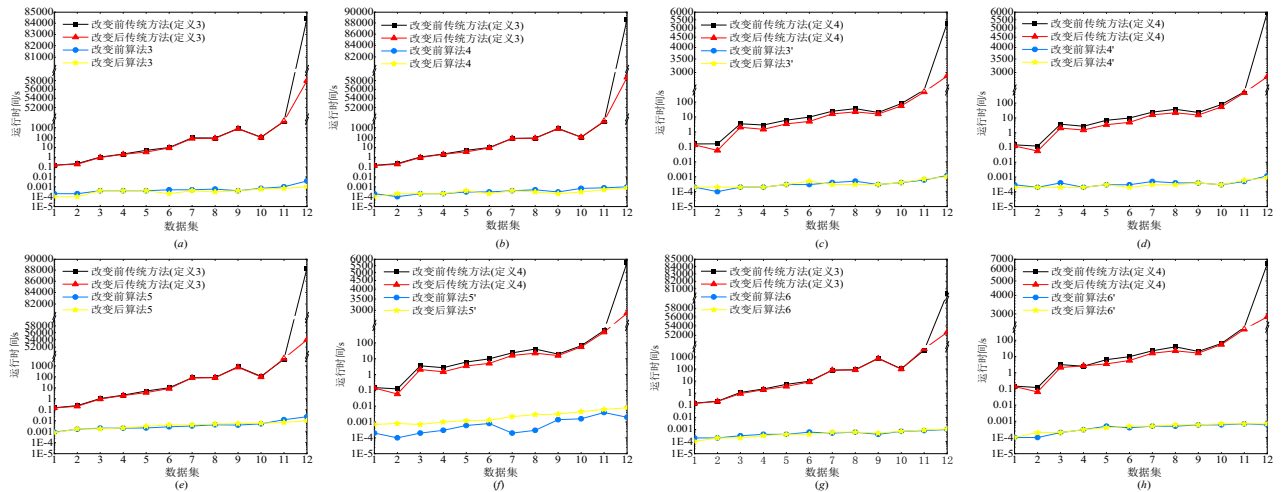


图3 数据集1~12对象结构信息类型改变前后的概念演化耗时对比

效性。

在本文工作的基础上,还可以进一步探讨以下问题:(1)当带对象结构信息形式背景中对象数增加或减少时,结构信息形式概念的演化情况;(2)带对象结构信息决策形式背景的概念认知学习问题;(3)由于粒计算在处理大数据方面有巨大的优势,如何融合粒计算与结构信息形式概念研究粒概念发现以提高计算效率也是有意义的。

参考文献

[1] FAWLEY W J, PIATETSKY-SHAPIRO G, MATHEUS C J. Knowledge discovery in databases: An overview[J]. AI Magazine, 1992, 13(3): 3-16.

[2] CERCONE N, TSUCHIYA M. Special issue on learning and discovery in knowledge-based databases-introduction [J]. IEEE Transactions on Knowledge and Data Engineer-

ing, 1993, 5(6): 901-902.

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

[4] 张文修, 梁怡, 吴伟志. 信息系统与知识发现[M]. 北京: 科学出版社, 2003.

ZHANG Wen-xiu, LEUNG Yee, WU Wei-zhi. Information Systems and Knowledge Discovery[M]. Beijing: Science Press, 2003. (in Chinese)

[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(11): 1478-1484.

HU Ke-yun, LU Yu-chang, SHI Chun-yi. An integrated mining approach for classification and association rule based on concept lattice[J]. Journal of Software, 2000, 11 (11): 1478-1484. (in Chinese)

[6] ZHANG Q, SHI C Y, NIU Z D, et al. HCBC: A hierarchi-

- cal case-based classifier integrated with conceptual clustering[J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 31(1): 152-165.
- [7] KAYTOUE M, KUZNETSOV S O, NAPOLI A, et al. Mining gene expression data with pattern structures in formal concept analysis[J]. Information Sciences, 2011, 181(10): 1989-2001.
- [8] TONELLA P. Using a concept lattice of decomposition slices for program understanding and impact analysis[J]. IEEE Transactions on Software Engineering, 2003, 29(6): 495-509.
- [9] 张文修, 魏玲, 祁建军. 概念格的属性约简理论与方法[J]. 中国科学E辑: 信息科学, 2005, 35(6): 628-639.  
ZHANG Wen-xiu, WEI Ling, QI Jian-jun. Attribute reduction theory and approach to concept lattice[J]. Science China Series E: Information Sciences, 2005, 35(6): 628-639. (in Chinese)
- [10] 魏玲, 曹丽, 祁建军, 等. 形式概念分析中的概念约简与概念特征[J]. 中国科学: 信息科学, 2020, 50(12): 1817-1833.  
WEI Ling, CAO Li, QI Jian-jun, et al. Concept reduction and concept characteristics in formal concept analysis[J]. Scientia Sinica Informationis, 2020, 50(12): 1817-1833. (in Chinese)
- [11] 李金海, 米允龙, 刘文奇. 概念的渐进式认知理论与方法[J]. 计算机学报, 2019, 42(10): 2233-2250.  
LI Jin-hai, MI Yun-long, LIU Wen-qi. Incremental cognition of concepts: Theories and methods[J]. Chinese Journal of Computers, 2019, 42(10): 2233-2250. (in Chinese)
- [12] 米允龙, 李金海, 刘文奇, 等. MapReduce 框架下的粒概念认知学习系统研究[J]. 电子学报, 2018, 46(2): 289-297.  
MI Yun-long, LI Jin-hai, LIU Wen-qi, et al. Research on granular concept cognitive learning system under MapReduce framework[J]. Acta Electronica Sinica, 2018, 46(2): 289-297. (in Chinese)
- [13] 智慧来. 面向异构数据分析的形式概念分析扩展模型[J]. 电子学报, 2013, 41(12): 2451-2455.  
ZHI Hui-lai. Extended model of formal concept analysis oriented for heterogeneous data analysis[J]. Acta Electronica Sinica, 2013, 41(12): 2451-2455. (in Chinese)
- [14] 张涛, 任宏雷, 洪文学, 等. 基于属性拓扑的可视化形式概念计算[J]. 电子学报, 2014, 42(5): 925-932.  
ZHANG Tao, REN Hong-lei, HONG Wen-xue, et al. The visualizing calculation of formal concept that based on the attribute topologies[J]. Acta Electronica Sinica, 2014, 42(5): 925-932. (in Chinese)
- [15] 翟岩慧, 李德玉, 曲开社. 决策蕴涵规范基[J]. 电子学报, 2015, 43(1): 18-23.  
ZHAI Yan-hui, LI De-yu, QU Kai-she. Canonical basis for decision implications[J]. Acta Electronica Sinica, 2015, 43(1): 18-23. (in Chinese)
- [16] QI J J, WEI L, CHEN Y P. Correlation analysis between objects and attributes[C]//Proceedings of 4th Rough Sets and Knowledge Technology. Berlin: Springer, 2009: 594-600.
- [17] RAGHAVAN V, HAFEZ A. Dynamic data mining[C]//Proceedings of 13th International Conference on Industrial Engineering Applications of AI & Expert Systems. New York: Springer, 2000: 220-229.
- [18] PEDRYCZ W, HOMENDA W. Building the fundamentals of granular computing: A principle of justifiable granularity[J]. Applied Soft Computing, 2013, 13(10): 4209-4218.
- [19] 张清华, 吕功勋, 陈玉洪, 等. 基于字符型属性值更新的动态三支决策模型[J]. 电子学报, 2019, 47(2): 344-350.  
ZHANG Qing-hua, LV Gong-xun, CHEN Yu-hong, et al. A dynamic three-way decision model based on the updating of character attribute values[J]. Acta Electronica Sinica, 2019, 47(2): 344-350. (in Chinese)
- [20] 马垣, 马文胜. 概念格多属性渐减式构造[J]. 软件学报, 2015, 26(12): 3162-3173.  
MA Yuan, MA Wen-sheng. Construction of multi-attributes decrement for concept lattice[J]. Journal of Software, 2015, 26(12): 3162-3173. (in Chinese)

#### 作者简介



**闫梦宇** 女, 1994年5月出生于山东省枣庄市. 现为昆明理工大学系统科学专业博士研究生. 主要研究方向为形式概念分析、复杂网络、概念认知学习等.

E-mail: yanmengyu2016@163.com



**李金海(通讯作者)** 男, 1984年1月出生于江西省上饶市. 2012年毕业于西安交通大学数学与统计学院. 现为昆明理工大学数据科学研究中心教授、博士生导师. 主要研究方向为概念格、粗糙集、粒计算、概念认知学习等.

E-mail: jhlixjtu@163.com