

数据库与信息处理

基于偏序压缩技术的频繁序列模式数据挖掘

程舒通^{1,2}, 徐从富¹, 但红卫¹

1. 浙江大学 计算机科学与技术学院, 杭州 310027

2. 杭州科技职业技术学院, 杭州 310022

收稿日期 修回日期 网络版发布日期 2008-1-11 接受日期

摘要 基于频繁序列模式的压缩技术旨在提高数据挖掘结果的可用性和可理解性, 从庞大的序列模式中发现有用的知识。分析了现有频繁模式压缩算法的不足, 提出了在模式聚类函数的基础上生成一个压缩的偏序 (Partial Order) 的算法, 实验结果显示该算法可以对频繁序列模式进行高效, 高质量的压缩, 可以得到数量更少、信息量更大的模式, 从而提高发现的频繁访问序列的兴趣性。

关键词 [数据挖掘](#) [频繁序列模式](#) [偏序](#) [压缩](#)

分类号

Partial order compresssion-oriented frequent sequential pattern datamining

CHENG Shu-tong^{1,2}, XU Cong-fu¹, DAN Hong-wei¹

1.College of Computer Science of Technology of Zhejiang University, Hangzhou 310027, China

2.Hangzhou Poly Technique College, Hangzhou 310022, China

Abstract

In this article, compression method based on frequent sequential pattern is to improve the usability and intelligibility of data mining, discover useful information from among colossal sequence. After analyzing the shortages of the existing algorithms for frequent pattern compression, we propose an algorithm for creating compressed partial order based on pattern clustering function. Experiment result shows the compression arithmetic has higher proficiency and quality, we can obtain much less number and much more information pattern. Thereby we could find out more interesting of visited frequent sequential than the normal algorithm.

Key words [data mining](#) [frequent sequential pattern](#) [partial order](#); [compression](#)

DOI:

通讯作者 程舒通 chengshutong@21cn.com

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF\(576KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [参考文献](#)

服务与反馈

▶ [把本文推荐给朋友](#)

▶ [加入我的书架](#)

▶ [加入引用管理器](#)

▶ [复制索引](#)

▶ [Email Alert](#)

▶ [文章反馈](#)

▶ [浏览反馈信息](#)

相关信息

▶ 本刊中 [包含“数据挖掘”的相关文章](#)

▶ 本文作者相关文章

· [程舒通](#)

·

· [徐从富](#)

·

· [但红卫](#)