

数据库、信号与信息处理

基于FS-tree的频繁模式挖掘算法

史旻昱, 马辉民, 唐述科

华中科技大学 管理学院, 武汉 430074

收稿日期 2008-3-6 修回日期 2008-5-23 网络版发布日期 2008-10-18 接受日期

摘要 关联规则挖掘是数据挖掘中的一个重要研究方向, 用于发现项集之间的关联性。FP-growth算法通过构造FP-tree产生频繁集, 由于其不生成候选集从而大大降低了搜索开销, 其缺点是占用大量的内存空间。基于FP-growth的算法思想, 提出基于FS-tree (频繁1-项子树) 的频繁模式挖掘算法, 通过将FP-tree拆分为多棵FS-tree, 使算法的空间复杂度明显减小。实验表明, 该算法是有效的。

关键词 [数据挖掘](#) [关联规则](#) [频繁1-项子树](#) [频繁模式](#)

分类号

Algorithm of frequent patterns mining based on FS-tree

SHI Min-yu, MA Hui-min, TANG Shu-ke

School of Management, Huazhang University of Science and Technology, Wuhan 430074, China

Abstract

Association rule mining which is used to find the correlation of items is an important research direction in data mining. FP-growth algorithm greatly reduces the search time without generating candidate itemsets by constructing FP-tree to find frequent itemsets. The drawback of it is taking a lot of memory space. Based on the thinking of FP-growth algorithm, an algorithm for mining frequent patterns based on FS-tree (1 item frequent sub-tree) is proposed. The algorithm reduces the space complexity significantly by splitting FP-tree into some FS-trees. The experiments show that the algorithm is effective.

Key words [data mining](#) [association rule](#) [FS-tree](#) [frequent patterns](#)

DOI: 10.3778/j.issn.1002-8331.2008.30.051

通讯作者 史旻昱 shiminyu@gmail.com

扩展功能

本文信息

- ▶ [Supporting info](#)
- ▶ [PDF\(549KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [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

- [史旻昱](#)
- [马辉民](#)
- [唐述科](#)