

数据库与信息处理

Apriori挖掘频繁项目集算法的改进

柴华昕, 王 勇

桂林电子科技大学 网络中心, 广西 桂林 541004

收稿日期 修回日期 网络版发布日期 2007-8-9 接受日期

摘要 针对Apriori算法的不足, 提出了一种新的优化算法Napriori。算法从优化产生2-项目集、事务压缩、项目压缩、优化连接等几个方面对Apriori算法进行优化, 将散列技术应用于产生1-项目集和2-项目集, 将压缩优化和连接优化应用于k-项目集。实验结果表明, Napriori算法运行速度比Apriori算法有了明显的提高。

关键词 [关联规则](#) [Apriori算法](#) [事务压缩](#) [散列结构](#)

分类号

Improvement of Apriori algorithm

CHAI Hua-xin, WANG Yong

NIC of Guilin University of Electronic Technology, Guilin, Guangxi 541004, China

Abstract

In this study, it proposes a new optimization algorithm called Napriori based on the insufficient of Apriori. Napriori algorithm presents optimizations on 2-items generation, transactions compression, items compression and join optimization. Napriori uses hash structure to generate 1-items and 2-items while compression and join optimization to generate k-items ($k > 3$). The performance study shows that Napriori is much faster than Apriori.

Key words [association rule](#) [Apriori algorithm](#) [transaction compression](#) [hash structure](#)

DOI:

通讯作者 柴华昕

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF\(1311KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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“关联规则”的相关文章](#)

▶ 本文作者相关文章

· [柴华昕](#)

· [王 勇](#)