

数据库、信息处理

TWCT-Stream: 数据流上的频繁模式挖掘算法

庄波¹, 刘希玉², 隆坤¹

1. 滨州学院 计算机科学技术系, 山东 滨州 256603

2. 山东师范大学 管理与经济学院, 济南 250014

收稿日期 2009-2-12 修回日期 2009-5-8 网络版发布日期 2009-7-9 接受日期

摘要 提出一种结合倾斜时间窗的TWCT树结构, 可以保存不同时间粒度下频繁模式的完全集, 并设计了其顺序更新和删除算法, 使其能够存储在外存, 从而有效地降低算法的内存空间需求。结合TWCT树结构特点, 提出了数据流上的频繁模式挖掘算法TWCT-Stream, 其模式生长的TWCT-Growth算法按字典顺序生成频繁模式, 以配合TWCT结构的顺序更新。实验证实算法的内存需求低于FP-Stream等同类算法。

关键词 [数据流挖掘](#) [频繁模式](#) [倾斜时间窗口](#)

分类号

TWCT-Stream: Algorithm for mining frequent patterns in data streams

庄波¹, 刘希玉², 隆坤¹

1. Department of Computer Science and Technology, Binzhou University, Binzhou, Shandong 256603, China

2. School of Management and Economics, Shandong Normal University, Jinan 250014, China

Abstract

A TWCT tree structure is proposed with tilted-time window framework embedded, which can maintain the complete set of frequent patterns at multiple time granularities. And this paper designs the sequential update and delete algorithms for the structure, which makes it can be saved in auxiliary storage in order to reduce the algorithms' requirements of the main memory effectively. Taking advantage of this characteristic, TWCT-Stream, a frequent pattern mining algorithm in data stream, is proposed. And its pattern growth algorithm, TWCT-Growth, generates frequent patterns in lexical order suitable for the sequential updating of TWCT structure. Experiments have proved its memory requirements lower than the same kind of algorithms like FP-Stream and so on.

Key words [data stream mining](#) [frequent pattern](#) [tilted-time window](#)

DOI: 10.3778/j.issn.1002-8331.2009.20.044

通讯作者 庄波 sdzhuangbo@126.com

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF\(533KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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“数据流挖掘”的相关文章](#)

▶ 本文作者相关文章

· [庄波](#)

· [刘希玉](#)

· [隆坤](#)