

- >> 首页
- >> 被收录信息
- >> 投稿须知
- >> 模板下载
- >> 信息发布
- >> 常见问题及解答
- >> 合作单位
- >> 产品介绍
- >> 编委会/董事会
- >> 关于我们
- >> 网上订阅
- >> 友情链接

友情链接

- >> 中国期刊网
- >> 万方数据资源库
- >> 台湾中文电子期刊
- >> 四川省计算应用研究中心
- >> 维普资讯网

数据流中基于滑动窗口的最大频繁项集挖掘算法*

Mining maximal frequent itemsets over data streams using sliding window

摘要点击: 38 全文下载: 68

[查看全文](#) [查看/发表评论](#) [下载PDF阅读器](#)

中文关键词: [数据流](#) [数据挖掘](#) [最大频繁项集](#) [滑动窗口](#) [位图](#)

英文关键词: [data stream](#) [data mining](#) [maximal frequent itemsets](#) [sliding window](#) [bitmap](#)

基金项目: 湖南省教育厅科研资助项目(08B040)

作者

[杨路明](#), [刘立新](#), [毛伊敏](#), [谢东](#)

单位

[\(中南大学 信息科学与工程学院, 长沙 410083\)](#)

中文摘要:

挖掘数据流中最大频繁项集是从数据流中获得信息的一种有效手段,是数据流挖掘研究的热点之一。结合数据流的特点,提出了一种新的基于滑动窗口的最大频繁项集挖掘算法。该算法用位图来存储数据流中流动的数据;采用直接覆盖的方法存储和更新数据流上的数据;在深度优先搜索挖掘最大频繁项集时,除采用经典的剪枝策略外,还提出了与父等价原理相对应的子等价剪枝策略;最后将挖掘结果存储在索引链表中以提高超集检测效率,进一步减少挖掘最大频繁项集的时间。理论分析和实验结果证实了该算法在时间和空间上的有效性。

英文摘要:

Mining maximal frequent itemsets over streaming data is one of the most important issues in mining data stream. This paper proposed an efficient algorithm to mine maximal frequent itemsets in sliding window. First, used bitmap to deal with the streaming data. Second, adopted depth first to find maximal frequent itemsets when mining. Moreover, besides typical pruning strategies, this paper developed a new pruning strategy corresponding to the parent equivalency pruning to prune. Third, used index structure to store the maximal frequent itemsets, which could speed up the speed of superset test. Theoretical analysis and experimental results show that the proposed method is efficient.

您是第2827010位访问者

主办单位: 四川省计算机研究院 单位地址: 成都市武侯区成科西路3号

服务热线: 028-85249567 传真: 028-85210177 邮编: 610041 Email: arocmag@163.com

蜀ICP备05005319号 本系统由北京勤云科技发展有限公司设计