

学术研究

## 数据流中基于矩阵的频繁项集挖掘

王磊+, 黄志球, 朱小栋, 沈国华, 程亮

南京航空航天大学 信息科学与技术学院, 南京 210016

收稿日期 修回日期 网络版发布日期 2008-6-20 接受日期

**摘要** 挖掘频繁项集是挖掘数据流的基本任务。许多近似算法能够有效地对数据流进行频繁项挖掘,但不能有效地控制内存资源消耗和挖掘运行时间。为了提高数据流频繁项集挖掘的时空效率,通过引入矩阵作为概要数据结构,提出了一种新的数据流频繁项集挖掘算法。最后通过实验证明了该算法的有效性。

**关键词** [数据流](#) [数据挖掘](#) [频繁模式](#) [矩阵](#)

分类号

## Mining frequent itemsets over data stream by matrix

WANG Lei+, HUANG Zhiqiu, ZHU Xiaodong, SHEN Guohua, CHENG Liang

College of Information Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

### Abstract

Mining frequent itemsets is a basic task of the data stream mining. Recently many approximate algorithms can mine frequent itemsets over data stream. However, these algorithms still can not efficiently reduce space and time cost. To improve the efficiency of mining frequent itemsets over data stream, matrix is imported as the synopsis data structure and a new algorithm of mining frequent itemsets is presented. Finally, experiments prove the efficiency of this algorithm.

**Key words** [data stream](#) [data mining](#) [frequent itemsets](#) [matrix](#)

DOI:

通讯作者 王磊 [ad82ad82@126.com](mailto:ad82ad82@126.com)

### 扩展功能

#### 本文信息

- ▶ [Supporting info](#)
- ▶ [PDF\(899KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [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](#)
- ▶ [浏览反馈信息](#)

#### 相关信息

- ▶ [本刊中 包含“数据流”的相关文章](#)
- ▶ [本文作者相关文章](#)

- [王磊](#)
- [黄志球](#)
- [朱小栋](#)
- [沈国华](#)
- [程亮](#)