

数据库、信号与信息处理

## 高激励项集的挖掘研究

余光柱<sup>1, 3</sup>, 刘旭辉<sup>2</sup>, 邵世煌<sup>1</sup>

1. 东华大学 信息科学与技术学院, 上海 201600

2. 长江大学 机械工程学院, 湖北 荆州 434000

3. 湖北警官学院 计算机系, 武汉 430034

收稿日期 2008-7-2 修回日期 2008-8-3 网络版发布日期 2009-12-4 接受日期

**摘要** 基于支持度的关联规则只能找出所有的频繁集, 无法找到那些非频繁但效用很高的项集; 基于效用的关联规则致力于发现所有高效用项集, 无法找到效用不高但支持度与效用的积很大的项集。为克服支持度与效用的不足, 提出了一种新的项集重要性的度量方法(即激励)及一种自下而上的挖掘高激励项集的算法HM-Two-Phase-Miner。激励集成了支持度与效用的优点, 能同时表达项集的语义特性与统计特性。HM-Two-Phase-Miner利用事务权重激励向下封闭特性进行减枝, 有效提高了算法的性能。

**关键词** [高激励项集](#) [关联规则](#) [支持度](#) [基于效用](#)

**分类号** [TP182](#)

## Study of high motivation itemsets mining

YU Guang-zhu<sup>1, 3</sup>, LIU Xu-hui<sup>2</sup>, SHAO Shi-huang<sup>1</sup>

1. College of Information Science and Technology, Donghua University, Shanghai 201600, China

2. College of Mechanical Engineering, Yangtze University, Jingzhou, Hubei 434000, China

3. Department of Computer Science, Hubei University of Police, Wuhan 430034, China

### Abstract

Algorithms for support-based association rules mining can only discover frequent itemsets, but can not discover the non-frequent itemsets with high utility values; Utility-based association rules mining aims at discovering high utility itemsets, without considering the itemsets whose utility values are not high but the product of the support and utility of the same itemset is very large. To solve the problem, a new measure is proposed, i.e., motivation, to measure the importance of an itemset and a down-top algorithm called HM-Two-Phase-Miner to discover high motivation itemsets. Motivation integrates the advantages of support and utility, and thus can reflect both the semantic significance and statistical significance of an itemset. In HM-Two-Phase-Miner algorithm, transaction-weighted motivation downward closure property is adopted to cut down the search space.

**Key words** [high motivation itemset](#) [association rule](#) [support](#) [utility-based](#)

DOI: 10.3778/j.issn.1002-8331.2009.33.041

通讯作者 余光柱 [guang216@126.com](mailto:guang216@126.com)

### 扩展功能

#### 本文信息

▶ [Supporting info](#)

▶ [PDF\(528KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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“高激励项集”的 相关文章](#)

▶ [本文作者相关文章](#)

· [余光柱](#)

· [刘旭辉](#)

· [邵世煌](#)