

博士论坛

## 一种基于划分的高效用长项集挖掘算法

余光柱<sup>1,4</sup>, 李克清<sup>2</sup>, 易先军<sup>3</sup>, 邵世煌<sup>1</sup>

1. 东华大学 信息学院, 上海 201600

2. 武汉大学 计算机学院, 武汉 430072

3. 北京大学 计算机软件与微电子学院, 北京 102600

4. 湖北省荆州市公安局 刑侦支队, 湖北 荆州 434000

收稿日期 修回日期 网络版发布日期 2007-9-29 接受日期

**摘要** 效用 (utility) 可弥补支持度在表现语义重要性方面的不足。现有的几种基于效用的关联规则挖掘算法都采用了类似Apriori自底向上的搜索方法, 不适合长模式的挖掘。提出了一种双向搜索高效用项集的模型及一种基于划分的inter-transaction算法。inter-transaction利用了长事务相交迅速变短的特性和新的减枝策略, 能同时输出项集的效用与支持度。实验表明, 该方法对蕴含长模式的高维数据库非常有效。

**关键词** [高效用长项集](#) [交集事务](#) [划分](#)

分类号

## Partition-based algorithm for mining high utility long itemsets

YU Guang-zhu<sup>1,4</sup>, LI Ke-qing<sup>2</sup>, YI Xian-jun<sup>3</sup>, SHAO Shi-huang<sup>1</sup>

1.College of Information Science and Technology,Donghua University,Shanghai 201600,China 2.College of Computer Science,Wuhan University,Wuhan 430072,China

3.School of Software and Microelectronics,Peking University,Beijing 102600,China 4.Criminal Investigation Brigade,Jingzhou Public Security Bureau,Jingzhou,Hubei 434000,China

### Abstract

Utility can be used to overcome the shortcoming of support.However,existing algorithms for utility mining adopt an Apriori-like bottom-up search and are inadequate on databases with long patterns.To solve the problem,we have proposed a hybrid model and an algorithm called inter-transaction to discover high utility itemsets from two directions.Inter-transaction utilizes a new pruning strategy and the characteristic that multiple long transactions usually have less common items,and it can output both the utility and support of an itemset.Experiments on synthetic data show that inter-transaction is very effective on high dimensional databases with long patterns.

**Key words** [long high utility itemset](#) [intersection transaction](#) [partition](#)

DOI:

通讯作者 余光柱 [E-mail: ygz@mail.dhu.edu.cn](mailto:ygz@mail.dhu.edu.cn)

### 扩展功能

#### 本文信息

▶ [Supporting info](#)

▶ [PDF\(1155KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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“高效用长项集”的 相关文章](#)

▶ 本文作者相关文章

· [余光柱](#)

· [李克清](#)

· [易先军](#)

· [邵世煌](#)