Abstract

Efficient discovery of frequent itemsets in large datasets is a crucial task of data mining. In recent years, several approaches have been proposed for generating high utility patterns, they arise the problems of producing a large number of candidate itemsets for high utility itemsets and probably degrades mining performance in terms of speed and space. Recently proposed compact tree structure, viz., UP-Tree, maintains the information of transactions and itemsets, facilitate the mining performance and avoid scanning original database repeatedly. In this paper, UP-Tree (Utility Pattern Tree) is adopted, which scans database only twice to obtain candidate items and manage them in an efficient data structured way. Applying UP-Tree to the UP-Growth takes more execution time for Phase II. Hence this paper presents modified algorithm aiming to reduce the execution time by effectively identifying high utility itemsets.

References

An Improved UP-Growth High Utility Itemset Mining

the 20th Int’l Conf. on Very Large Data Bases, pp. 487-499, 1994.


- Frequent itemset mining implementations repository, http://fimi.cs.helsinki.fi/


**Index Terms**

Computer Science

Information Sciences

**Keywords**

High utility itemsets  Transaction Weight Utilization  Utility Mining  Discarding