Abstract

Utility-based data mining is a new research area interested in all types of utility factors in data mining processes [1]. The basic meaning of utility is the quantity sold, interest, importance & profitability of items to the users. Utility of items in a transaction database consists of two aspects: 1. The importance of distinct or unique items, which is called external utility. 2. The importance of the items in the transaction, which is called as internal utility. Mining high utility itemsets from the databases is not an easy task. Pruning search space for high utility itemset mining is difficult because a superset of a low utility itemset may be a high utility itemset. Existing studies [2,4,9] applied overestimated methods to facilitate the mining performance of utility mining. In these methods, first we will get potential high utility itemsets, and then an additional database scan is performed for identifying their utilities. However, the existing methods often generate a huge candidate itemsets and the mining performance is degraded consequently. In this paper we proposed Eliminating Unusual Itemset by Eliminating item set which is low utility item set to reduce search space. Proposed methods not only reduce the number of candidate itemsets, but also significantly increase the performance of the mining process.
Mining High Utility Itemsets from Large Dynamic Dataset by Eliminating Unusual Items

References

- J. Han, J. Wang, Y. Lu and P. Tzvetkov, "Mining Top-k Frequent Closed Patterns without Minimum Support," Proc. of ICDM, 2002.


Index Terms

Computer Science
Information Sciences

Keywords
High Utility Mining Frequent Itemset Mining Eliminating Unusual Itemset Profit Quantity