Overview of Itemset Utility Mining and its Applications

Abstract

An emerging topic in the field of data mining is Utility Mining. The main objective of Utility Mining is to identify the itemsets with highest utilities, by considering profit, quantity, cost or other user preferences. Mining High Utility itemsets from a transaction database is to find itemsets that have utility above a user-specified threshold. Itemset Utility Mining is an extension of Frequent Itemset mining, which discovers itemsets that occur frequently. In many real-life applications, high-utility itemsets consist of rare items. Rare itemsets provide useful information in different decision-making domains such as business transactions, medical, security, fraudulent transactions, retail communities. For example, in a supermarket, customers purchase microwave ovens or frying pans rarely as compared to bread, washing powder, soap. But the former transactions yield more profit for the supermarket. Similarly, the high-profit rare itemsets are found to be very useful in many application areas. For example, in medical application, the rare combination of symptoms can provide useful insights for doctors [21]. A retail business may be interested in identifying its most valuable customers i.e. who contribute a major fraction of
overall company profit[10]. Several researches about itemset utility mining were proposed. In this paper, a literature survey of various algorithms for high utility rare itemset mining has been presented.

Reference

- R. Chan, Q. Yang, Y. D. Shen, “Mining High utility Itemsets”, in Proc. of the 3rd IEEE Intel. Conf. on Data Mining (ICDM), 2003.
- V. S. Tseng, C.J. Chu, T. Liang, “Efficient Mining of Temporal High Utility Itemsets from Data streams”, Proceedings of Second International Workshop on Utility-Based Data Mining, August 20, 2006
- G.C.Lan, T.P.Hong and V.S. Tseng, “A Novel Algorithm for Mining Rare-Utility Itemsets in a Multi-Database Environment”

Index Terms

Computer Science  Data Mining

Key words

Utility Mining

High-utility itemsets

Rare itemsets

Frequent Itemset mining