Abstract

Mining of frequent patterns is a basic problem in data mining applications. Frequent Itemset Mining is considered to be an important research oriented task in data mining, due to its large applicability in real world applications. In this paper, a new Maximal Frequent Itemset mining algorithm with effective pruning mechanism is proposed. The proposed algorithm takes vertical tidset representation of the database and removes all the non-maximal frequent item-sets to get exact set of MFI directly. Pruning is done for both search space reduction and minimizing the number of frequency computations. It works efficiently when the number of item-sets and tid-sets are more. The proposed approach has been compared with Mafia algorithm for mushroom dataset and the results shows that the proposed algorithm performs effectively and generates frequent patterns faster. In order to understand the algorithm easily, an example is provided in detail.

References

- Roberto Bayardo, "Efficiently mining long patterns from databases", in ACM
SIGMOD Conference 1998.
- R. Agarwal, C. Aggarwal and V. Prasad, "A tree projection algorithm for
- K. Gouda and M. J. Zaki, "Efficiently Mining Maximal Frequent Itemsets", in Proc.
of the IEEE Int. Conference on Data Mining, San Jose, 2001.
- Gosta Grahne and Jianfei Zhu, "Efficiently using prefix-trees in Mining Frequent
Itemsets", in Proc. of the IEEE ICDM Workshop on Frequent Itemset Mining
- Burdick, D., M. Calimlim and J. Gehrke, "MAFIA: A maximal frequent itemset
algorithm for transactional databases", In International Conference on Data Engineering,
pp: 443 – 452, April 2001, doi = 10. 1. 1. 100. 6805
- J. Han, J. Pei, and Y. Yin. "Mining frequent patterns without candidate
discovery of association rules", Advances in Knowledge Discovery and Data Mining,
- V. Ganti, J. E. Gehrke, and R. Ramakrishnan, "DEMON: Mining and Monitoring
Evolving Data", ICDE 2000: 439-448
- Aggarwal, C. C. and P. S. Yu, "Mining large itemsets for association rules", in
Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 1998, pp:
- Aggarwal C. C and P. S. Yu, "Online generation of association rules", in
- Park J. S, M. S. Chen, P. S. Yu, "An Effective Hash Based Algorithm for Mining
ISSN: 0163-5808.
- Dunkel B. and N. Soparkar, "Data Organization and access for efficient data
mining", in the proceedings of the 15th International Conference on Data Engineering, pp:
- R. Agrawal, T. Imielienski and A. Swami, "Mining association rules between sets
ACM SIGMOD Conference on Management of Data, Pages 207-216, Newyork, 1993, ACM
Press.

Index Terms

Computer Science

Data Mining
Keywords
Data Mining  Frequent Itemset Mining  Maximal Frequent Itemset Mining