Abstract

An Important Problem in Data Mining in Various Fields like Medicine, Telecommunications and World Wide Web is Discovering Patterns. Frequent patterns mining is the focused research topic in association rule analysis. Apriori algorithm is a classical algorithm of association rule mining. Lots of algorithms for mining association rules and their mutations are proposed on basis of Apriori Algorithm. Most of the previous studies adopt Apriori-like algorithms which generate-and-test candidates and improving algorithm strategy and structure but no one concentrate on the structure of database. A simple approach is if we implement in Transposed database then result is very fast. Recently, different works proposed a new way to mine patterns in transposed databases where a database with thousands of attributes but only tens of objects. In this case, mining the transposed database runs through a smaller search space. In this paper, we systematically explore the search space of frequent patterns mining and
represent database in transposed form. We developed an algorithm (termed DFPMT—A Dynamic Approach for Frequent Patterns Mining Using Transposition of Database) for mining frequent patterns which are based on Apriori algorithm and used Dynamic function for Longest Common Subsequence [1]. The main distinguishing factors among the proposed schemes is the database stores in transposed form and in each iteration database is filter/reduce by generating LCS of transaction id for each pattern. Our solutions provide faster result. A quantitative exploration of these tradeoffs is conducted through an extensive experimental study on synthetic and real-life data sets.

Reference

- J. Han, Research challenges for data mining in science and engineering. In NGDM 2007.
- A fast APRIORI implementation Ferenc Bodon∗ Informatics Laboratory, Computer and Automation Research Institute, Hungarian Academy of Sciences H-1111 Budapest, L´agym´anyosi u. 11, Hungary
- Improving Frequent Patterns Mining by LFP XU Yusheng, MA Zhixin, CHEN Xiaoyun, Li Lian School of Information Science and Engineering Lanzhou University Lanzhou, China, 730000 e-mail:{xuyusheng, mazhx, chenxy, lil}@lzu.edu.cn Tharam S. Dillon School of Information System Curtin University Perth, Australia
- Finding Longest Increasing and Common Subsequences in Streaming Data David Liben-Nowell_ y dln@theory.lcs.mit.edu Erik Vee_ z env@cs.washington.edu An Zhu_ x anzhu@cs.stanford.edu November 26, 2003
- F. Bodon and L. R´onyai. Trie: an alternative data structure for data mining algorithms. to
An Implementation of Frequent Pattern Mining Algorithm using Dynamic Function


- Efficient Mining of Weighted Frequent Pattern Over Data Streams. Farhan Ahmed, Tanbeer 2009

Index Terms

Computer Science Data Mining

Key words

Longest Common Subsequence Transposition of Database

Frequent Pattern mining