Efficient Algorithms for Mining Frequent Patterns

Ming-Chen Liu, Chien-Chung Hsu

Abstract

Data mining is a very important database research issue. Especially, the generation of frequent patterns in large databases has been widely studied. Most of the studies take the Apriori-based approach, which has great effort in the generation of candidate frequent patterns and needs multiple database accesses. The FP-Tree-based approaches have been proposed to avoid the generation of candidate set and scan transaction database only twice, but they work with more complicated data structure. Recently, a frequent pattern list (FPL) algorithm, using a simple linear list to store all transactions, was proposed to improve the FP-Tree algorithm. However, FPL algorithm still needs to scan database twice. In this paper, an efficient frequent pattern generation algorithm, called FPLI, was proposed to improve the FPL algorithm. FPLI scans the database only once and uses a simple linear list to store all transactions like FPL. By performing simple operations on the list, we can discover the frequent patterns quickly. It is also not necessary for FPLI to rescan database and reconstruct data structure when transaction database is updated or minimum support is varied. Experimental results show that the FPLI algorithm has much better performance than the FPL algorithm.

Keywords: Data Mining, Frequent Pattern, Candidate Itemset

Table of Contents

Chapter 1: Introduction
Chapter 2: Related Study
  2.1 Apriori Algorithm
  2.2 DHP Algorithm
  2.3 DIC Algorithm
  2.4 Random Sampling Algorithm
  2.5 PINCER-Search Algorithm
  2.6 FP-Tree Algorithm
  2.7 FPL Algorithm
Chapter 3: Fast High-Frequency Itemset Mining Algorithm
  3.1 Problem Definition and Exploration
  3.2 FPLI Algorithm
    3.2.1 Lexicon Description
    3.2.2 Algorithm Analysis and Description
    3.2.3 Complexity Analysis
Chapter 4: Application for Minimum Support Change and Database Modification
  4.1 Minimum Support Change Algorithm
  4.2 Addition of Transaction Algorithm
  4.3 Deletion of Transaction Algorithm
Chapter 5: Experiment and Result Analysis
  5.1 Experiment Test Example Explanation
  5.2 Experiment Result Analysis and Performance Evaluation
    5.2.1 FPLI Algorithm and FPL Algorithm Performance Comparison
    5.2.2 Minimum Support Change Algorithm and FPL Algorithm Performance Comparison
    5.2.3 Addition of Transaction Algorithm and FPL Algorithm Performance Comparison
    5.2.4 Deletion of Transaction Algorithm and FPL Algorithm Performance Comparison
Chapter 6: Conclusion and Future Prospect

References
