URN etd-0407109-183710-787
Author ANNIE CHANG
Author's Email Address No Public.
Statistics This thesis had been viewed 41 times. Download 0 times.
Department
Year 2005
Semester 2
Degree Master
Type of Document Master's Thesis
Language zh-TW.Big5 Chinese
Title Mining Association Rules Algorithm with Positive Correlation and Closed Itemsets
Date of Defense 0000-00-00
Page Count 77
Keyword
positive correlation itemsets
closed itemsets
data mining
association rule

Abstract
The correlation among the itemsets of data mining in the transactions intensified the effect of decision supporting system in the organization. The algorithm of data mining technology plays the key roles for the effectiveness of data mining and the utilities of the resourses. One of the most defective problem of the previous researches for the correlative data mining algorithm has low effectiveness for system filtering. Too many regulations which have been generated by the decision supporting systems cause the poor efficiency. To present a simple and effective correlation algorithm is the main objective of this research.

A Positive Correlation and Closed Itemsets by Phase(PCP) algorithm was presented in this reaseach to modify Gradation Reduction Approaches(GRA) algorithm by adding concept of closed itemsets and positive correlation itemsets. Although GRA algorithm has gradually reduced a great deal of works in the transactions of database, and can reduce a great number of infrequent itemsets, but the decision makers are confused by too many frequent itemsets and association rules which are generated by GRA algorithm when they want to make decisions.

The PCP algorithm use the concept of closed itemsets and positive correlation itemsets to reduce the number of association rules and the mining results will be meaningful. The size of the databases in the real world is always greater than the size of the memory. In order to solve this problem, we propose a modifying algorithm - PCP-M(Positive Correlation and Closed Itemsets by Phase - Modified Version); it divides a large database into many sub-databases and mines association rules from those sub-databases. The PCP-M algorithm only scans database four times and will not be affect by the length of frequent itemsets. The PCP-M algorithm avoids wasting a lot of I/O time and increases the efficiency and the practicability in application.
Advisory Committee - advisor

Files No Any Full Text File.
Date of Submission 2006-06-27