

Sorted Compressed Tree: An Improve Method of Frequent Patterns Mining without Support Constraint

邱創楷, 曾秋蓉

Computer Science & Information Engineering

Computer Science and Informatics

judyCRT@chu.edu.tw

Abstract

Several algorithms have been proposed for association rule mining, such as Apriori and FP Growth. In these algorithms, a minimum support should be decided for mining large itemsets. However, it is usually the case that several minimum supports should be used for repeated mining to find the satisfied collection of association rules. To cope with this problem, several algorithms were proposed to allow the minimum support to be adjusted without rebuilding the whole data structure for frequent pattern mining. The Compressed and Arranged Transaction Sequences tree (CATS tree) algorithm is one of them. Nevertheless, CATS Tree builds its tree structure dynamically, so that the mining process is complex and tedious. In this paper, we present an improved algorithm called the Sorted Compressed tree (SC tree). By pre-sorting the datasets, the tree structure can be built statically. Moreover, association rules can be mined in a bottom-up style instead of bi-directional in CATS tree and recursive in FP Growth. Hence, the cost of association rule mining is reduced. From preliminary experimental results, SC tree is not only more efficient but is also space saving.

Keyword : association rule mining; without support constrain; CSTS tree; SC Tree