A new method for mining maximal frequent itemsets

ABSTRACT

In this paper, we propose a new method for mining maximal frequent itemsets. Our method introduces an efficient database encoding technique, a novel tree structure called PC_Tree and also PC_Miner algorithm. The database encoding technique utilizes Prime number characteristics and transforms each transaction into a positive integer that has all properties of its items. The PC_Tree is a simple tree structure but yet powerful to capture whole of transactions by one database scan. The PC_Miner algorithm traverses the PC_Tree and builds the gcd (greatest common divisor) set of its nodes to mine maximal frequent itemsets. Experiments verify the efficiency and advantages of the proposed method.