A numerical method for frequent pattern mining

ABSTRACT

Frequent pattern mining is one of the active research themes in data mining. It plays an important role in all data mining tasks such as clustering, classification, prediction, and association analysis. Identifying all frequent patterns is the most time consuming process due to a massive number of patterns generated. A reasonable solution is identifying maximal frequent patterns which form the smallest representative set of patterns to generate all frequent patterns. In this paper, an efficient numerical method for mining frequent patterns is proposed. This method is based on prime number characteristics to generate all frequent patterns by using maximal frequent ones. There are two new properties introduced in this method; a novel tree structure called PC_Tree and PC_Miner algorithm. The PC_Tree is a simple tree structure but yet capable to capture the whole of transactions information with an efficient data transformation technique that utilizes the prime number theory. The PC_Miner algorithm traverses the PC_Tree by using an efficient pruning technique. The experimental results verify the compactness and the efficiency of mining shown by the proposed method.