Summary by Deeksha Ramakrishna

Efficient Infrequent Itemset Mining Using Depth-First and Top-Down Lattice Traversal

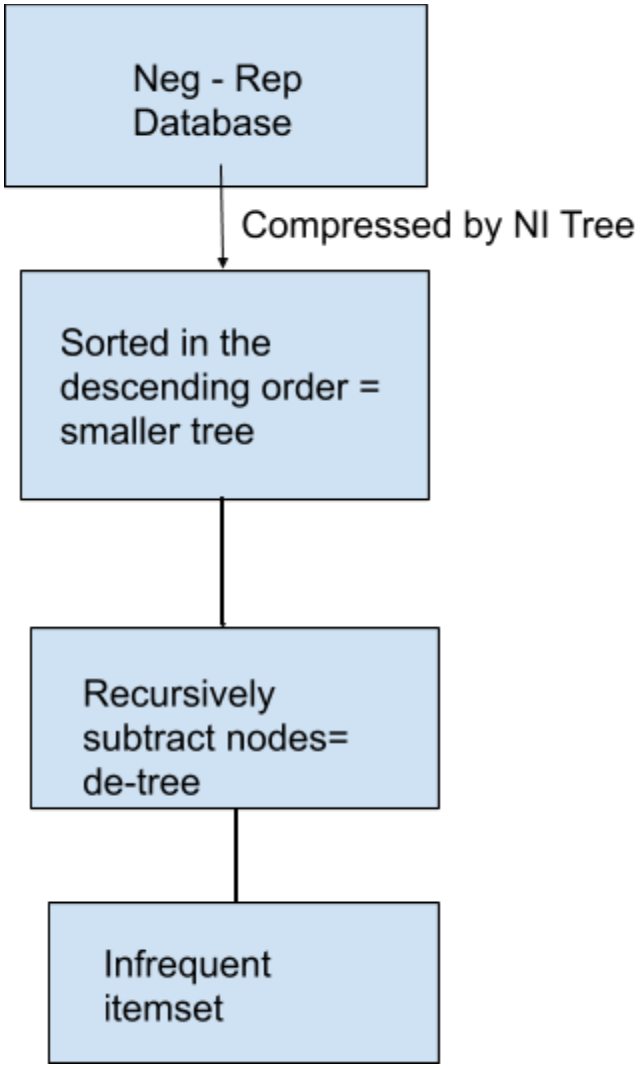
This paper proposes an efficient rare pattern extraction algorithm, which is capable of extracting the complete set of rare patterns using a top-down traversal strategy. A negative item tree is employed to accelerate the mining process. Pattern growth paradigm is used and therefore avoids expensive candidate generation.

Problem Statement : The main purpose is to address the problem of exhaustive time intensive consideration to find out the rare patterns and also the huge number of candidates generated in the mining process using the bottom up traversal in Apriori algorithm.

Proposed Approach : Negative Infrequent Itemset tree miner (NIIMiner) is the algorithm proposed. The general idea here is to build the negative itemset tree in the same way as the well known FP-tree by scanning the whole neg-rep database. Negative items are sorted in descending order which leads to a smaller tree.

Infrequent itemsets are extracted by recursively subtracting (excluding) nodes from the NI-tree. The new tree after subtraction, called deducted tree (de-tree), is also an NI-tree. Items that have been subtracted so far are stored in the root node.

Negative items are excluded one by one recursively. The whole process terminates until the joint support is lower than the given threshold.



Experiment Evaluation ：The NIIMiner is compared with the Rarity algorithm which is the state of the art top-down Apriori like infrequent itemset mining approach. The proposed algorithm is significantly faster compared to the Rarity algorithm under most of the settings. The only drawback here is it does not perform well in case of under large minimum support value and small maximum itemset size settings. In summary this has the ability to extract the complete set of rare patterns, by utilizing the negative representations of rare itemsets to frequent itemsets and also on the sparse data the Rarity algorithm works faster.