

Finding frequent closed itemsets with an extended version of the Eclat algorithm

Laszlo Szathmary

University of Debrecen, Faculty of Informatics, Department of IT
H-4002 Debrecen, Pf. 400, Hungary
szathmary.laszlo@inf.unideb.hu

Abstract

Apriori is the most well-known algorithm for finding frequent itemsets (FIs) in a dataset. For generating interesting association rules, we also need the so-called frequent closed itemsets (FCIs) that form a subset of FIs. *Apriori* has a simple extension called *Apriori-Close* that can filter FCIs among FIs. However, it is known that vertical itemset mining algorithms outperform the *Apriori*-like levelwise algorithms. *Eclat* is another well-known vertical miner that can produce the same output as *Apriori*, i.e. it also finds the FIs in a dataset. Here we propose an extension of *Eclat*, called *Eclat-Close* that can filter FCIs among FIs. This way *Eclat-Close* can be used as an alternative of *Apriori-Close*. Experimental results show that *Eclat-Close* performs much better than *Apriori-Close*, especially on dense, highly-correlated datasets.

Keywords: data mining, frequent itemsets, association rules, algorithms

MSC: 97R40