Hypergraph Transversals in Association Rule Mining

Cristina Tîrnăucă

Cristina Tîrnăucă (`cristina.tirnauca@unican.es`)
Universidad de Cantabria

Abstract. Association rule mining is a popular and well researched method for discovering useful relations between items in a dataset. For example, in a market basket dataset, one might conclude after analyzing the data that *bread, sausages → mustard*, that is, "most of the customers that bought bread and sausages also bought mustard". Association rule generation is often divided into two separate steps: minimal support is applied to find all frequent itemsets, which are subsequently used to form rules. In practice, the number of frequent itemsets is often huge, so condensed representations, with sizes that can be several orders of magnitude smaller than the size of frequent set collections, are the norm.

In particular, *minimal generators* are widely used in order to prune non-redundant rules. Unfortunately, finding minimal generators is equivalent to finding minimal transversals in a certain hypergraph (the relation between minimal hypergraph transversals and data mining has already been noted in [1], for example), and whether all minimal transversals can be computed in output-polynomial time is still an open problem.

The aim of this presentation is to offer a brief introduction of the above mentioned concepts and the relevant literature, hopefully leading to fruitful collaborations that may shed some light on the above mentioned problem.

References

[1] Gunopulos, D.; Khardon, R.; Mannila, H.; Toivonen, H. Data Mining, Hypergraph Transversals, and Machine Learning. *Proc. PODS* (1997), 209–216.