

Origin based Association Rule Mining using multiple MASP tree^{☆,☆☆}

C.V. Radhakrishnan^{a,1,*}, K. Bazargan^{a,b,2}, S. Pepping^{c,1,3,**}

^a*River Valley Technologies, SJP Building, Cotton Hills, Trivandrum, Kerala, India 695014*

^b*River Valley Technologies, 9, Browns Court, Kennford, Exeter, United Kingdom*

^c*Central Application Management, Elsevier, Radarweg 29, 1043 NX
Amsterdam, Netherlands*

Abstract

Association rule learning is a rule-based machine learning method for discovering interesting relations between variables in large databases.

Keywords: data-mining, Association Rule Mining, frequent-itemset mining

1. Introduction

Association rule mining is a rule-based machine learning procedure to find interesting patterns in the transaction database based on individual and conditional frequencies. In the traditional approach, two steps are involved in generating rules. First, generate all frequent itemsets and pruned non-frequent ones and then in the second stage rules are derived from those frequent itemsets. An association rule e.g. {bread, milk} \Rightarrow {butter} in market basket analysis means if one purchase bread and milk together it is highly likely that they will also buy butter. Apart from market basket analysis, association rule mining is useful in intrusion detection, bioinformatics, and many other applications.

In 2014 Omer M. Soyal [1] proposed a new approach to extract mostly associated sequential patterns (MASPs) using less computational resources in terms of time and memory while generating a long sequence of patterns that have the highest co-occurrence. This approach may produce different outcomes if we

[☆]This document is a collaborative effort.

^{☆☆}The second title footnote which is a longer longer than the first one and with an intention to fill in up more than one line while formatting.

*Corresponding author

**Principal corresponding author

Email addresses: cvr@river-valley.com (C.V. Radhakrishnan),
kaveh@river-valley.com (K. Bazargan)

URL: <http://www.elsevier.com> (S. Pepping)

¹This is the specimen author footnote.

²Another author footnote, but a little more longer.

³Yet another author footnote. Indeed, you can have any number of author footnotes.

change the order of items in transactions. We propose an approach which is order independent. An association rule of the form $A \implies B$ must satisfy the threshold support and threshold confidence i.e. probability of occurrence of A and B together must surpass threshold support, and the probability of occurrence of B in transactions containing A must be greater than or equal to threshold confidence. It means, to calculate support and confidence, it is required to traverse complete transaction database. To generate all rules containing a particular item x it is reasonable to ignore all transactions(for calculating support and confidence) that come before the transaction in which that particular item appears for the first time. Embedding these two changes to the Omer M. Soyal [1] approach is the basis of our research.

References

- [1] O. M. Soyal, Association rule mining with mostly associated sequential patterns, Elsevier.