

# VisAR : A New Technique for Visualizing Mined Association Rules

Kesaraporn Techapichetvanich and Amitava Datta

School of Computer Science & Software Engineering,  
University of Western Australia, Perth, WA 6009, Australia  
{kes, datta}@csse.uwa.edu.au

**Abstract.** Many business organizations generate a huge amount of transaction data. Association rule mining is a powerful analysis tool to extract the useful meanings and associations from large databases and many automated systems have been developed for mining association rules. However, most of these systems usually mine many association rules from large databases and it is not easy for a user to extract meaningful rules. Visualization has become an important tool in the data mining process for extracting meaningful knowledge and information from large data sets. Though there are several techniques for visualizing mined association rules, most of these techniques visualize the entire set of discovered association rules on a single screen. Such a dense display can overwhelm analysts and reduce their capability of interpretation. In this paper we present a novel technique called *VisAR* for visualizing mined association rules. VisAR consists of four major stages for visualizing mined association rules. These stages include *managing association rules*, *filtering association rules of interest*, *visualizing selected association rules*, and *interacting with the visualization process*. Our technique allows an analyst to view only a particular subset of association rules which contain selected items of interest. VisAR is able to display not only many-to-one but also many-to-many association rules. Moreover, our technique can overcome problems of screen clutter and occlusion.

**Keywords:** Visualization, Visual Exploration, Association Rules, Market Basket.

## 1 Introduction

Many business organizations generate a huge amount of transaction data. Association rule mining is a powerful analysis tool to extract the useful meanings and associations for such data. Information visualization plays a major role to enhance the capability of analysts for analyzing large amounts of data. There is some existing research on visualization of association rules [1, 2, 3]. However, one of the drawbacks of these visualization techniques is that visualizing all mined association rules on a single screen reduces the interpreting and understanding ability of analysts. Typically, association rules generated by mining algorithms are difficult for users to understand due to their large number. Visualization allows users to visually analyze and understand the generated association rules.

This paper presents a new technique called *VisAR* for visualizing association rules derived from the mining process. We focus on reducing the complexity of visualizing

large number of association rules on a single screen so that users are able to effectively understand and interpret information from a large number of association rules. We have also designed our system to eliminate occlusion. Users can explore association rules through their specified items of interest using an interactive visualization scheme. Our technique is able to represent both many-to-one and many-to-many association rules. Moreover, our technique allows users to select items appearing in the antecedents of association rules so that the users can view only the association rules containing their items of interest.

The rest of the paper is organized as follows. We discuss some previous work in Section 2, we explain the VisAR system in Section 3, the advantages of the VisAR system are presented in Section 4 and finally we conclude with some comments in Section 5.

## 2 Previous Work

### 2.1 Association Rules

An association rule [4] is formally described as a rule of the type  $A \Rightarrow B$  where  $A$  is an item set called *antecedent*, *body*, or *left-hand side* (LHS) and  $B$  is an item set called *consequent*, *head*, or *right-hand side* (RHS). Each item set consists of items from a transactional database. Items existing in the antecedent are not in the consequent. In other words, an association rule is  $A \Rightarrow B$  where  $A, B \subset I$  and  $A \cap B = \phi$ .  $I = \{i_1, i_2, \dots, i_n\}$  is a set of items in the transaction database where  $i_j$ ,  $1 \leq j \leq n$ , is an item in the database that may appear in a transaction. The two common measures of interestingness are *support* and *confidence*. The support of a rule is defined as the percentage of frequency with which all items in the rule appear together. The confidence of the rule is the ratio of frequency of items in both antecedent and consequent (frequency of  $A$  and  $B$ ) to frequency of items in antecedent appearing together. The probability of both support and confidence is,  $\text{Support}(A \Rightarrow B) = P(A \cup B)$ ,  $\text{Confidence}(A \Rightarrow B) = P(B|A)$ .

A term, *frequent itemset* [5] or *large itemset* [6] is used to define item sets whose number of co-appearances in the database is greater than a user specified support. In other words, these items are frequently purchased together and their occurrence is higher than the specified minimum support.

### 2.2 Related Work

Visualization techniques are integrated into data mining to help users understand the data as well as discover associations and pattern in the data. Various methodologies have been developed for visualizing association rules that are generated by an automated data mining algorithm. Prior research for visualizing association rules can be categorized into three main groups: *Table-based*, *Matrix-based*, and *Graph-based*.

First, Table-based techniques are the most common and traditional approaches to represent association rules in the form of a table. In general, the columns of a rule table represent the items, the antecedents and consequents, the support, and the confidence of association rules. Each row represents an association rule. Some examples of Table-based techniques can be found in SAS Enterprise Miner [7] and DBMiner [5].