

A NEW FRAMEWORK OF MINING ASSOCIATION RULES WITH TIME-WINDOWS ON REAL-TIME TRANSACTION DATABASE

YIYONG XIAO¹, RENQIAN ZHANG² AND IKOU KAKU³

¹School of Reliability and System Engineering

²School of Economics and Management

Beihang University

No. 37, Xueyuan Road, Haidian District, Beijing 100191, P. R. China

{ xiaoyiyong; zhangrenqian }@buaa.edu.cn

³Department of Management Science and Engineering

Faculty of Systems Science and Technology

Akita Prefectural University

Yulihonjo, Akita 015-0055, Japan

ikou_kaku@akita-pu.ac.jp

Received October 2009; revised February 2010

ABSTRACT. *This paper dedicates efforts to discover the part-time association rules in real-time transactional database by extending the traditional minsup-minconf based framework to a new one – the minsup, minconf and minwin based framework. We propose a more general form for association rule, i.e., the Association Rule with Time-Windows (ARTW), to properly integrate the temporal association rules together with the normal ones. New notions like Frequent Itemset with Time-Windows (FITW) are also defined, and an Apriori-like algorithm, named TW-Apriori, is developed to fast generate the FITWs. Computational experiments are conducted on two datasets – a synthetic dataset and a real database. Both experiments show that large number of ARTWs ignored previously can be discovered under the new framework; many of them are even very strong rules and valuable for market decisions. The efficiency of the proposed TW-Apriori algorithm is also proven feasible since it can finish the calculation within one minute and the length of the calculation time is nearly proportional to the number of ARTWs found.*

Keywords: Data mining, Association rule, Association rule with time-window, Temporal association rule, Periodic association rule

1. **Introduction.** Association Rule Mining (ARM) aims at discovering interesting relationships among a set of items by finding itemset that frequently appeared together in transactions. Agrawal et al. [1] proposed firstly an ARM approach to generate all the association rules that meet the user-specified minimum support (*minsup*) and minimum confidence (*minconf*). Agrawal and Srikant [2] decomposed the ARM problem into two sub-problems: (1) Find all itemsets that have support above *minsup*, namely, the frequent itemsets; (2) Generate association rules from frequent itemsets that satisfy *minconf*. Since the first problem is the most time-consuming part of ARM algorithms, many subsequent researches were mainly focused on finding more efficient algorithms for discovering frequent itemsets. A typical approach called Apriori/AprioriTid presented by Agrawal and Srikant [2] can fast generate frequent itemsets by using the knowledge that all subsets of a frequent itemset are also frequent. This knowledge considerably simplifies the discovery of frequent itemsets. Since there are thousands of researches on ARM, typical reviews can be found for example in Han and Micheline [3] and Zhang and Zhang [4].