

# A Weighted Load-Balancing Parallel Apriori Algorithm for Association Rule Mining

游坤明, Jia-Ling Zhou

Computer Science & Information Engineering

Computer Science and Informatics

yu@chu.edu.tw

## Abstract

Because of the exponential growth in worldwide information, companies have to deal with an ever growing amount of digital information. One of the most important challenges for data mining is quickly and correctly finding the relationship between data. The Apriori algorithm is the most popular technique in association rules mining; however, when applying this method, a database has to be scanned many times and many candidate itemsets are generated. Parallel computing is an effective strategy for accelerating the mining process. In this paper, the Weighted Distributed Parallel Apriori algorithm (WDPA) is presented as a solution to this problem. In the proposed method, metadata are stored in TID forms, thus only a single scan to the database is needed. The TID counts are also taken into consideration, and therefore better load-balancing as well as reducing idle time for processors can be achieved. According to the experimental results, WDPA outperforms other algorithms while having lower minimum support.

Keyword : Data Mining, Association Rule, Apriori, Cyclic Distribution, Loading Balance