Parallel Frequent Patterns Mining Algorithm on GPU

周嘉奕,游坤明,吳秉璋
Computer Science & Information Engineering
Computer Science and Informatics
yu@chu.edu.tw

Abstract

Extraction of frequent patterns from a transactional database is a fundamental task in data mining. Its applications include association rules, time series, etc. The Apriori approach is a commonly used generate-and-test approach to obtain frequent patterns from a database with a given threshold. Many parallel and distributed methods have been proposed for frequent pattern mining (FPM) to reduce computation time. However, most of them require a Cluster system or Grid system. In this study, a graphic processing unit (GPU) was used to perform FPM with a GPU-FPM to speed-up the process. Because of GPU hardware delimitations, a compact data structure was designed to store an entire database on GPU. In addition, MemPack and CLProgram template classes were also designed. Two datasets with different conditions were used to verify the performance of GPU-FPM. The experimental results showed that the speed-up ratio of GPU-FPM can achieve 14.857 with 16 times of threads.

Keyword：frequent pattern mining, parallel processing, graphic processing unit (GPU), OpenCL