

An OpenCL Candidate Slicing Frequent Pattern Mining Algorithm on Graphic Processing Units

林哲瑜, 游坤明, 歐陽雯, Jiayi Zhou

Computer Science & Information Engineering

Computer Science and Informatics

ouyang@chu.edu.tw

Abstract

Frequent pattern mining (FPM) is important in data mining field with Apriori algorithm to be one of the commonly used approaches to solve it. However, Apriori algorithm encounters an issue that the computation time increases dramatically when data size increases and when the threshold is small. Many parallel algorithms have been proposed to speed up the computation using computer clusters or grid systems. GPUs have also been applied on FPM with only few adopting OpenCL although OpenCL has the advantage of being platform independent. Thus, the aim of this research is to develop efficient parallel Apriori strategy using GPU and OpenCL. Our novel method, Candidate Slicing Frequent Pattern Mining (CSFPM) algorithm, improves over the previous method by slicing candidate information to better balance the load between processing units. This strategy is proved to be more efficient according to our experiments. For example, CSFPM is at most 2.6 times faster than the previous method. Therefore, CSFPM is an efficient parallel Apriori algorithm which can reduce computation time and improve overall performance.

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