

発表概要

分散環境での並列グラフマイニングにおける タスク中断処理による冗長探索削減

奥野 伸吾^{1,†1} 平石 拓^{2,a)} 中島 浩² 八杉 昌宏³ 瀬々 潤⁴

2017年7月28日発表

本発表では、各頂点がアイテムの集合を持つようなグラフからユーザが指定した閾値以上の数の共通アイテム集合を持つ連部結合グラフを全抽出するグラフマイニングアルゴリズム COPINE の、タスク並列言語 Tascell を用いた既存実装の改善を提案する。COPINE は、左深さ優先の探索木走査中に得た知識を蓄えるテーブルを利用した枝刈りにより不要な探索を回避するが、並列探索では、枝刈りの実行が他ワーカによる走査開始後になることがあるため、無視できない量の無駄な探索が生じてしまう。これを解決するため、脱出対象となる try ブロック内の全並列タスクを自動的に中断する機能を持つ例外処理を用いて無駄な探索を速やかに中断させる手法を提案し、すでに共有メモリ環境ではその効果を確認している。本研究では、この手法を分散環境にも適用した。この結果、16 ノード × 16 ワーカを用いた性能評価で探索空間削減効果が得られることが確認できた。

Presentation Abstract

Reducing Redundant Search in Parallel Graph Mining Using Task Abortion in Distributed Memory Environments

SHINGO OKUNO^{1,†1} TASUKU HIRAISHI^{2,a)} HIROSHI NAKASHIMA² MASAHIRO YASUGI³ JUN SESE⁴

Presented: July 28, 2017

This presentation proposes an improvement of an existing parallel implementation of the graph mining algorithm called COPINE using the Tascell task parallel language. COPINE extracts all connected subgraphs with common itemsets, of which the size is not less than a given threshold, from a graph and itemsets associated with vertices of the graph. This algorithm avoids unnecessary searches using a pruning mechanism that depends on the knowledge acquired during the left- and depth-first traversal of its search tree. In the parallel search, a problem arises that a worker may prune the subtree after another worker has started to traverse it, resulting in a large number of redundant searches. Such redundancy can be significantly reduced by letting a worker abort such redundant traversals as soon as possible. This abortion can be implemented efficiently using exception handling whereby all parallel tasks running in a try block with an exception are automatically aborted. We have already proposed this mechanism and realized performance improvements in shared memory environments. In this research, we applied it to distributed memory environments. As a result, we successfully reduced search space in performance evaluations using 16 nodes × 16 workers.

This is the abstract of an unrefereed presentation, and it should not preclude subsequent publication.

¹ 京都大学大学院情報学研究所
Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
² 京都大学学術情報メディアセンター
Academic Center for Computing and Media Studies, Kyoto University, Kyoto 606-8501, Japan
³ 九州工業大学大学院情報工学研究院
Department of Artificial Intelligence, Kyushu Institute of Technology, Iizuka, Fukuoka 820-8502, Japan

⁴ 産業技術総合研究所
National Institute of Advanced Industrial Science and Technology, Koto, Tokyo 135-0064, Japan
^{†1} 現在、株式会社富士通研究所
Presently with Fujitsu Laboratories Ltd.
^{a)} tasuku@media.kyoto-u.ac.jp