

論文 / 著書情報  
Article / Book Information

題目(和文)	大規模動的グラフ解析の性能最適化手法
Title(English)	Performance Optimization of Large-Scale Dynamic Graph Analytics
著者(和文)	金刺宏樹
Author(English)	Hiroki Kanezashi
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第11222号, 授与年月日:2019年6月30日, 学位の種別:課程博士, 審査員:松岡 聡,増原 英彦,遠藤 敏夫,脇田 建,額田 彰
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第11222号, Conferred date:2019/6/30, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

(博士課程)

## 論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	金刺 宏樹		
論文審査 審査員		氏名	職名		氏名	職名
	主査	松岡 聡	特任教授	審査員	額田 彰	准教授
	審査員	増原 英彦	教授			
		遠藤 敏夫	教授			
		脇田 建	准教授			

### 論文審査の要旨 (2000 字程度)

本論文は「Performance Optimization of Large-Scale Dynamic Graph Analytics (大規模動的グラフ解析の性能最適化手法)」と題し、ソーシャルネットワークなどの大規模動的グラフデータ上での解析アルゴリズムとエージェントベースシミュレーションの性能最適化を目的としている。この最適化を達成するために、動的グラフ解析において発生する性能上の課題に対処するアルゴリズムを提案し、実社会データによる評価によって高速化が得られていることを示している。本論文は英文により全7章から構成されている。

第1章では、大規模動的グラフを用いた解析とシミュレーションの意義について述べ、その際に発生する性能上の問題点を指摘している。これらの問題を解決するために、グラフコミュニティ検出、グラフパターンマッチング、道路交通シミュレーション、それぞれのアプリケーションにおいて提案した性能最適化アルゴリズムの概要と重要性を述べている。

第2章では、第1章で述べた大規模動的グラフ解析の問題点と背景について、実際のデータセットを提示しながら詳細を述べている。さらに、本論文で提案するアルゴリズムを理解するために必要な、コミュニティ検出とグラフパターンマッチングの代表的なアルゴリズム、エージェントベースの道路交通シミュレーションのフレームワークについて説明している。

第3章では、並列分散環境上でのエージェントベースの道路交通シミュレーションの性能最適化手法について説明している。最初に、大規模な道路交通シミュレーションフレームワークとそれを実行するための並列分散環境の必要性について述べ、計算ノード間通信のオーバーヘッド及び計算負荷の不均衡などそれらに伴う性能上の問題提起を行っている。その上でこれらの問題を軽減するために、車両エージェントの動的割り当てによる負荷の均等化、ノード間同期の頻度削減手法を提案している。この手法をダブリン市の道路ネットワークと、通勤車両のトリップデータという2つのケースで評価し、それぞれ最大25%、35%の高速化を示している。

第4章では、インクリメンタルなグラフコミュニティ検出アルゴリズム IncrementalDEMON を提案している。既存のコミュニティ検出アルゴリズム DEMON をベースラインとし、内部の主要な処理をそれぞれインクリメンタル化することで、動的なグラフにおいても余分な再計算を最低限に抑えている。また、性能上のボトルネックとなっている微細なコミュニティ同士の比較処理についても最適化を施し、18万~244万エッジのソーシャルネットワークにおいて、1回のインクリメンタル実行が DEMON アルゴリズムと比較して平均1/101の時間で済むことを示している。

第5章では、より高速で正確なサブグラフパターンを検出するためのインクリメンタルなグラフパターンマッチングアルゴリズムと、そのパラメータを動的に調整する IGPM-PEM を提案している。最初に実社会データ解析において必要なグラフパターンマッチングの条件について述べ、それらを満たすためのインクリメンタルアルゴリズム IGPM と、IGPM で再計算するサブグラフの範囲を強化学習によって動的に調整する機構 PEM について説明している。これらの提案手法により、387万エッジのソーシャルネットワークにおいて、1回あたりのインクリメンタル実行が既存のインクリメンタルでないアルゴリズム G-Ray と比較して平均1/14.8の時間で済むことが示されている。

第6章では、第3、4、5章でそれぞれ提案した性能最適化手法について性能面、精度面双方について

て評価を行い、有用性、実用性について考察している。性能面に関しては、各最適化手法の適用前後での実行時間の定式化を行い、性能向上の効果について述べている。精度面では、各手法で近似アルゴリズムを使用していることから、最適化手法を適用した前後で発生する誤差の評価を行っている。

第7章では、本論文の一連の研究を総括し、結論と貢献を述べるとともに、さらなる大規模化と高速化に向けた今後の発展性について述べている。

以上のように、本研究は実社会の動的グラフデータを使用した解析、シミュレーションの性能に関する問題提起を行い、解決のための最適化手法を提案検証し、その有効性を確認しており、理学上貢献するところ大である。よって、本論文は博士（理学）の学位論文として十分価値があるものと認める。

注意：「論文審査の要旨及び審査員」は、東工大リサーチポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。