

「ビッグデータ統合利活用のための次世代基盤技術の創出・体系化」
平成 26 年度採択研究代表者

H27 年度
実績報告書

加藤 直樹

関西学院大学工学部
教授

ビッグデータ時代に向けた革新的アルゴリズム基盤

§ 1. 研究実施体制

(A)「劣線形アルゴリズム」グループ

- ① 研究代表者:加藤 直樹 (関西学院大学, 教授)
- ② 研究項目
 - ・ビッグデータ向け定数時間アルゴリズムの実用化と効率化
 - ・漸進型アルゴリズムの開発
 - ・線形時間アルゴリズムの開発
 - ・組合せ剛性理論によるタンパク質の機能解明
 - ・避難計画問題
 - ・革新的アルゴリズム基盤の構築

(B)「劣線形データ構造」グループ

- ① 主たる共同研究者:渋谷 哲朗 (東京大学医科学研究所, 准教授)
- ② 研究項目
 - ・情報論的アプローチによる劣線形データ構造に関する研究
 - ・列挙論的アプローチによる劣線形データ構造に関する研究
 - ・実応用アプローチによる劣線形データ構造に関する研究

(C)「劣線形モデリング」グループ

- ① 主たる共同研究者:田中 和之 (東北大学大学院情報学研究科, 教授)
- ② 研究項目
 - ・統計力学的粗視化アプローチによる劣線形モデリング
 - ・計算理論と統計的近似理論の融合による大規模システム上での高効率計算アルゴリズムの開発

§ 2. 研究実施の概要

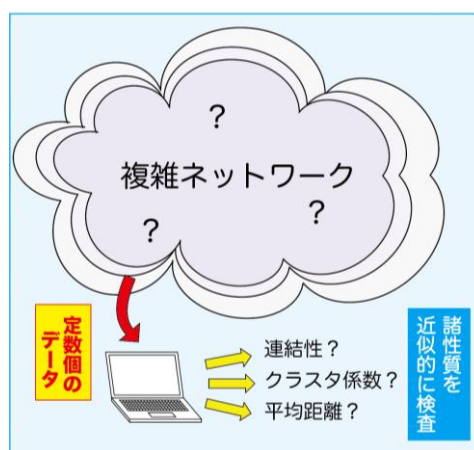
研究実施概要

今世紀に新たに注目を浴びているビッグデータは、そのデータ量の膨大さ故に、その基礎となるアルゴリズムのモデルは根本的な変革を迫られている。例えば、これまで速いと考えられてきたアルゴリズムが、ペタスケールのビッグデータに対して計算資源や実行時間などの点で大きな困難が伴い、少なくとも線形、場合によっては劣線形時間や定数時間アルゴリズムが求められる。

そのために、本研究では、ビッグデータ時代に向けた新しい計算パラダイムとして「劣線形時間パラダイム」を提唱し、その中で、ビッグデータ用のアルゴリズムとデータ構造、およびモデリング技法を提案する。本研究は、劣線形時間アルゴリズムグループ、劣線形データ構造グループ、劣線形モデリングの3つの研究グループから構成されている。以下では、本年度の成果として学術的、および社会的にも注目を浴びた研究の概要を紹介する。

劣線形アルゴリズムグループの伊藤は、複雑ネットワークのモデル化である多重グラフのクラス HSF (Hierarchical-Scale-Free) に対して、このクラスは超有限性を持つことを証明した。この結果により、多くの複雑ネットワークに対し、少なくとも理論的には「すべての性質の検査がグラフのサイズに無関係な定数時間でできる」ことになり、ビッググラフデータに対する大きな進展である(図 1)。

また、同グループの瀧澤は、避難計画問題への効率的アルゴリズムの適用を担当しており、主に現在は世界的に大規模で複雑な地下街である大阪梅田地下街を対象としている。本年度は同協議会が平成 28 年 3 月に公表した「大阪駅周辺地区 地下空間浸水対策計画 Ver.1」に、昨年度実施した地下街の避難シミュレーションの知見が反映された。また、民放テレビ 1 件、新聞 5 件に本研究活動の一端が紹介されるなど、既に成果が社会に還元され始めている。同じく、同グループの Sljoka は組合せ剛性理論を用いたたんぱく質機能解析を担当しているが、26 年度に発表した、タウたんぱく質の機能解析に関する論文に、アメリカの製薬会社に関心を持ち、アルツハイマー、パーキンソン病、老人性痴ほう症の治療薬の開発に向けて、共同研究が始められている。これらの病気に対しては、これまで、治療薬が存在しなかったこともあり、成果が期待される。



複雑ネットワークの定数時間検査

図 1

劣線形データ構造グループの坂本・山際は、劣線形データ構造の実現のため、ストリームデータを圧縮可能なロスレスデータ圧縮アルゴリズムを開発し、ナノ秒のオーダーでデータの途切れなく高速に圧縮・復号化が可能になる圧縮技術を構築した。処理性能の優位に加え、ハードウェアリソースと圧縮率を調整することができるため、小さなリソースであっても、適応的にハードウェア実装が可能である。従来と比べ、100倍から1000倍のメモリのみしか用いないオンライン・アルゴリズムの実現によって、FPGA (field-programmable gate array) の極めて小さなメモリ上での実装が可能となり、実際に実装した基盤上でも、画面の転送速度を大幅に向上することに成功した(図 2)。

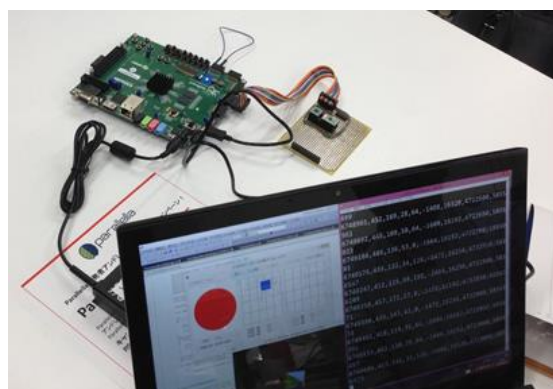
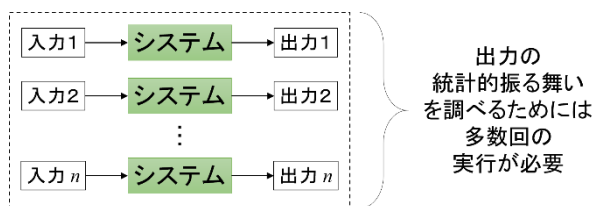


図 2: FPGA によるリアルタイム圧縮アルゴリズムの実装

劣線形モデリンググループでは、劣線形モデリングにおいて重要な役割を果たす**確率伝搬法**の高速化を実現した。確率伝搬法において、様々な入力データに対する出力の統計的な振る舞いを調べることで、構築したシステムの健全性などの多くのことを検証することができる。

この検証に対して従来は、様々な入力パラメータに対する出力の統計を取ることで達成してきた。したがって、正確な統計結果を得るためにはそれ相応の回数 of アルゴリズム処理を実際に行わなくてはならなくなる。

本グループは、確率伝搬法におけるモデルパラメータの変化に対する出力変化を高速に調べるアルゴリズムを統計力学理論を応用して開発した。その方法では、本来ならば膨大な回数 of アルゴリズム処理を実行しなくては得られない統計結果を、たった 1 回のアルゴリズム処理で高速に得ることができる (図 3)。このアルゴリズムは情報理論と統計力学理論を組み合わせたものであり、その双方に知見をもつ当グループならではの成果である。



提案法: 実質 1 回の実行で出力の統計性を知ることができる

図 3

代表的な原著論文

1. Tasuku Soma and Yuichi Yoshida, “Non-convex Compressed Sensing with the Sum-of-Squares Method”, Proceedings of 27th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 570-579, 2016
2. Shinichi Yamagiwa, Koichi Marumo, and Hiroshi Sakamoto, “Stream-based Lossless Data Compression Hardware using Adaptive Frequency Table Management”, LNCS, Springer, vol. 9495, pp. 1-14, 2015.

3. Muneki Yasuda, Shun Kataoka, and Kazuyuki Tanaka, “Statistical analysis of loopy belief propagation in random fields”, *Physical Review E*, vol. 92, Issue 4, pp. 042120, 2015