

JST/ERATO

河原林巨大グラフプロジェクト Kawarabayashi Large Graph Project

News Letter vol.01

ニュースレター vol.01

Contents

Overview プロジェクト概要

Message 研究統括の挨拶

Voices 各研究グループの紹介

Event イベント

Outreach アウトリーチ

2015.01

アルゴリズムが世界を変える
— *Algorithms Change the World.*

ご挨拶

研究統括 河原林健一

2年間を振り返って

本プロジェクトも発足して2年たちました。正式な立ち上げは2012年10月1日でしたが、私にとっては、ERATO^{*1}の内定をいただいた8月上旬からいただいた2年と数ヶ月たったことになります。

現在まで一番大変だった時期は、今考えれば内定をいただいた直後から発足までだったように思います。ERATOは巨大なプロジェクトですので、立ち上げの段階で、どのように人を集めるかがとても重要な要素になります。特に本プロジェクトのように大型実験装置を必要としないプロジェクトの場合、適材適所に若い研究員を配置することが、成否に直結します。この作業におそらく2012年8～10月の、ほぼすべてを費やしたように思います。それと同時に研究計画と予算計画が必要となってくるわけですが、こちらに関しては当時、JST研究プロジェクト推進部の南裕一氏にほぼ丸投げ状態でほとんどやっていただきました。JSTは、研究者へのRespectを最大限に示し、かつ研究者の事務的負担は最低限にしようとするシステムを持っています。このシステムのおかげで、本プロジェクト発足後2ヶ月の準備期間を乗り越えられたように思います(その代わり、研究成果を挙げなければすべては、研究統括の責任になります。No way outであることはたしかですが)。

ところで、ERATOを研究機関に持ってくるのは大変なこと!のようで、例えば、北海道大学の湊真一先生^{*2}は学長から直々に激励の言葉を受けたそうです。東京大学でERATOをもらうのはそれほど珍しくないことなので、そこまで大きなセレモニーは行われませんが、その他の研究機関では大学の学長の訪問やら、地方では地元の新聞で特集が組まれたりしているそうです。わたしの場合はこのようなことはありませんでしたが。

また文部科学省にとってもERATO研究統括は特別な存在で、20年以上前にERATOの研究統括になられた山本喜久先生^{*3}によると、当時のERATO研究統括は、文部大臣に直接研究計画を紹介する機会があったそうです。ちなみに山本先生は、江田五月大臣(当時)に研究紹介をしたそうです。私のときは、そのような機会がありませんでしたが、本プロジェクト発足時の文科大臣は田中眞紀子氏でした(もちろん民主党政権です)!ですからもしこのようなチャンスがあれば、それはそれで楽しかったように思います。

2年間の研究成果

研究成果はいろいろな尺度ではかることができるので、一概に表に出ている成果だけで判断することは難しいのですが、過去2年の本プロジェクトの「表に出ている」研究業績は悪くはないと思っています。とくに計算機科学分野のTIER 1(当該分野トップ会議)に本プロジェクトの40本以上の論文が採択されたこと。それからアルゴリズム分野のトップ会議であるSODA^{*4}において、直近3回のうち2回(2013、2015年)で本プロジェクトの論文がベストペーパー賞をとったことを考えると、過去2年の研究業績はまずまずで

あったと考えるべきでしょう。

本プロジェクトを始めるに当たって、最初考えたのは日本のコンピューターサイエンスのVISIBILITYをどのように上げていくか?ということでした。日本の中でも世界的な研究成果を上げている研究者は何人かいらっしゃいます。しかしとにかく層が薄い。ERATO発足時にいろいろなコンピューターサイエンスの分野を調査しましたが、どの分野も日本のVISIBILITYは、トップ会議においては2～3%でした(なかには5%近くの分野、限りなく0%に近い分野もありましたが、ほぼすべての分野で2～3%でした)。したがって本プロジェクトの最初の数年間は、まず世界に通用する研究者を育成する場にすることを考えました。

この目標を達成するためには、すでに完成された研究者より学位をとる前後の若手研究者にチャレンジしてもらったほうが、変に日本固有の価値観にとまっていけないので、はるかに成功する確率が高いと考えています(実際、優秀な大学院生も現在は非常に多いです)。現在の本プロジェクトの陣容は、自分の考えどおりの若い編成になり、彼らにはできる限りの研究環境を与えようと思っています。

今後の展望

最初の2年間が終わりました。世界で戦えそうな若手研究者も5～10人程度、本プロジェクトから生まれてきました。今後の本プロジェクトの展望について、以下のように考えています。

1 若い力を結集させた新たなプロジェクト

今後は世界的にもインパクトのある仕事をやっていかないとはいけないと思います。世界で戦っていくためには、生産性を高めるだけでなく、コンピューターサイエンス全体に貢献できるような仕事も考えていかないとはいけません。今後は、世界の価値観を共有した若手研究者と一緒に新たなコンピューターサイエンスの方向性を探っていきたいと思っています。

2 人材の更なる底上げ

本プロジェクトも、おかげさまで業界内での知名度も上がってきました。それと同時に、研究員のリクルートも本格的に始まっています。今後は研究員の出入りが頻繁に起こるのではと思っています。研究員が本プロジェクトでの仕事が評価され、他の研究機関に「栄転」できるのであれば、本プロジェクトとしては、素直に喜ぶべきであると思います。そしてまた新しい人材を育てればよいように思います。

3 日本のコンピューターサイエンスの基礎研究による底上げ

本プロジェクトは、基礎研究を中心にした人材の供給と日本のコンピューターサイエンスの底上げに貢献したいと思っています。同じ志がある研究機関や企業には、喜んで人材を供給したいと思っています。

- *1. ERATO Exploratory Research for Advanced Technologyの略。JST(科学技術振興機構)のプログラムの一つで、卓越したリーダーの元、独創性に富んだ探索型基礎研究を提案する。
- *2. 湊真一先生 ERATO湊離散構造処理プロジェクト(2009-2014)研究統括
- *3. 山本喜久先生 現JST/革新的研究開発推進プログラム(ImPACT)プログラム・マネージャー、前FIRST(最先端研究開発支援プログラム)「量子情報処理プロジェクト」代表研究者
- *4. SODA ACM-SIAM Symposium on Discrete Mathematicsの略。ACM(Association for Computing Machinery)とSIAM(Society for Industrial and Applied Mathematics)が開催する応用数学の会議。

河原林健一 Ken-ichi KAWARABAYASHI

国立情報学研究所情報学プリンシプル研究系教授/ビッグデータ数理国際研究センター長。専門分野は離散数学におけるグラフ彩色問題、グラフ構造理論とアルゴリズム、ネットワークフローとバス問題。

主な受賞に、文部科学省文部科学大臣表彰若手科学者賞[2006]、日本IBM日本IBM科学賞コンピューター・サイエンス部門[2008]、井上科学振興財団2009年井上リサーチアワード[2009]、公益財団法人船井情報科学振興財団船井学術賞船井哲良特別賞(アルゴリズムのグラフマイナー理論の研究とその応用)[2011]、日本学術振興会賞、日本学士院奨励賞、ACM-SIAM Symposium on Discrete Algorithms(SODA12) Best Paper Award[2013]などがある。

データから学び法則を導く グラフマイニング&ウェブ&AI グループ



グループリーダー

河原林健一 Ken-ichi KAWARABAYASHI
前ページ参照



サブリーダー

林浩平 Kohei HAYASHI
国立情報学研究所ビッグデータ数理国際研究センター特任助教

〈グラフマイニング&ウェブ&AI〉グループは、機械学習やデータマイニング*の技術を用いて現実世界に存在する巨大グラフを扱うことを目的として設立されました。

設立1年目の2013年度の時点で、専任研究員は林、大輪の2名、RAに関してはBowe、小野寺の2名と、他のグループと比較しても小規模で、カバーできている研究領域も限られていました。そこで、まず解くべき重要な問題をメンバー間で共有するため、外部から関連分野の研究者を積極的に関連分野の研究者を招待し、計20名の方に講演を行ってもらい、議論を繰り返しながら研究の方向性を固めていきました。

ここで創出された方向性の具体例としては、Twitterに代表される巨大かつ時変なグラフのストリーム処理アルゴリズムの開発が挙げられます。詳細については未発表のため省略しますが、この技術を用いることで、Twitterにおける日本ユーザの全つぶやきを処理し、その時々話題となっているトピックをリアルタイムで抽出することが可能となります。

2年目の2014年度では、専任研究員として波多野、馬場、菌部の3名が、またRAとして小西、Chelly、斉藤の3名が新たに着任し、カバーする研究領域も機械学習・クラウドソーシング・SATソルバ・インターネット広告・SNS解析・簡潔データ構造・情報検索と、大幅に拡大しました。また企業との共同研究にも着手し、より応用に直結する分野でのプロジェクトも現在進行中です。

このように各研究員・RAによってさまざまな研究がなされています。来年度はこれらの成果が対外的に発表できるよう、これを第一の目標として集中的に取り組みます。また現在連携中のいくつかの企業においては実データの取得やサービスの提案が非常にやりやすい環境となっています。これを活かし、アカデミックな成果と並行して実サービスに向けたシステムの開発・運用にも挑戦します。(林浩平)

データマイニング Data Mining

統計学、パターン認識、人工知能等のデータ解析の技法を大量のデータに網羅的に適用することで知識を取り出す技術。



グループセミナーの様子(国立情報学研究所)

実世界ネットワークを科学する 複雑ネットワーク&地図グラフ グループ



グループリーダー

吉田悠一 Yuichi YOSHIDA
国立情報学研究所情報学プリンシプル研究系特任助教/ビッグデータ数理国際研究センター副センター長

プロジェクト発足から今までを振り返ってみますと、個人的にはこれまでと異なる分野に挑戦した時期だったと言えます。学生の間は理論計算機科学という数学的な分野で研究をしていたわけですが、プロジェクトが始まってからはデータベース・AI*・データマイニングと言った様々な応用系の分野に論文を出すことが出来ました。実態としてはグラフアルゴリズムの高速化が中心で、それ専用の国際会議が無い為に結果的に色んな分野に分散しているという状況です。

ただグラフアルゴリズムの高速化とは一言では言いますが、「我々だからこそ出来ることをする」ことを心がけているつもりです。例えば、ヒューリスティクスで速くなりましたと言うだけではなく性能を理論的に解析したり、何故高速になったのかを実ネットワークの性質から理解しようとしていたりしています。前者は私が理論計算機科学をやっていることが役に立っていますし、後者は本グループにネットワーク科学を専門とするメンバーが揃っているからこそ出来ているのだと思います。また逆にアルゴリズムの挙動から実ネットワークの性質を明らかにしようということもしています。ある特定の分野の専門家というのは沢山いますが、本グループの様に一つのことを複数の観点から同時に見られる環境というのはなかなか無いので、大変幸せな環境にいと実感しています。

とは言え、折角これだけのメンバーが居るのに、グラフの研究ばかりで終わるのも勿体ないと思っています。グラフ以外にも、アルゴリズムも洗練されていなければ、その性質も分かっていない実データは沢山あります。例えば行列や時系列データはその最たる例でしょう。また個人的にはこれらの実データの研究を通じて、“Beyond Worst Case Analysis”、つまり最悪時計算量より現実的な計算量の解析は出来ないかという理論的な問い、に貢献できないか考えているところです。(吉田悠一)



グループセミナーの様子(国立情報学研究所)

人工知能 AI: artificial intelligence

人工的にコンピュータ上などで人間同様の知能を実現させようという試み。機械学習とよばれる技術を用いて、人間が自然に行っている学習能力と同等の機能をコンピュータで実現しようとする技術。

ネットワークで理論と応用をつなぐ グラフ・ネットワークにおける 理論と最適化グループ



グループリーダー

垣村尚徳 Naonori KAKIMURA

東京大学大学院総合文化研究科附属国際環境学
教育機構特任講師/国立情報学研究所客員
講師

〈グラフ・ネットワークにおける理論と最適化〉グループの目的は、組合せ最適化*やグラフ理論の先進的な手法を、通信ネットワークなどの分野に応用することです。本プロジェクトの他のグループは「ウェブグラフ」「複雑ネットワーク」などグループ名に対象とするネットワークを冠していますが、本グループは対象を限定せず、理論を幅広い応用分野に生かしていくことを目的としています。

グループのメンバーには理論計算機科学や離散数学の分野で活躍している30歳前後の若手研究者が数多くいます。メンバーの専門は、近似アルゴリズム、グラフ理論、離散構造、計算代数、数値計算など多種多様であり、それぞれが持つ理論的な知見を融合して実用上の問題を解決していくことが期待されています。

これまでに、たとえば、前原ら(VLDB*)はページランクの高速計算手法を提案しています。これは、木分解と呼ばれるグラフ理論の重要な概念を用いることで行列を分解し、その分解に応じてガウスの消去法・クリロフ部分空間法などの線形計算手法を使い分けることで計算を効率化しており、グラフ理論・数値計算など多様なバックグラウンドがあったからこそ産まれたものと言えます。

小林・大槻(INFOCOM*)は、通信ネットワークの耐故障性モデルに対して、効率的なアルゴリズムを設計しています。実用的なアルゴリズムを提案するだけでなく、古典的な最大フロー最小カット定理に対応するような最大最小定理を示しており、理論的にも深い結果になっています。

このようにメンバーの活躍のおかげでいくつかの成果が生み出されてきています。また、グループとしては応用分野の勉強会を毎週行っています。理論系分野とは観点が異なることが多いですが、扱っている数理モデルや手法は似ている点もあり、我々が貢献できる部分も多いことが分かってきました。今後も、自分たちの理論的バックグラウンドを生かしつつ、新しい分野を切り拓いていきたいと思っています。(垣村尚徳)

組合せ最適化 Combinatorial Optimization

有限個の解の中から最も良い解を見つけるための方法論。ネットワーク構造を持つ問題を扱うことが多い。有名な問題に巡回セールスマン問題などがある。

VLDB、INFOCOMはこの分野でのトップ会議。



グループミーティングの様子(東京大学)

ネットワークの諸問題は 素敵なアルゴリズムで解決 ネットワーク・アルゴリズム グループ



グループリーダー

徳山豪 Takeshi TOKUYAMA

東北大学大学院情報科学研究科教授/国立情報学研究所客員教授

本プロジェクトの4つの研究グループのうち、唯一東京の外(東北大学)に拠点を置いているのが〈ネットワーク・アルゴリズム〉グループです。当グループでは、ネットワークの構成やネットワーク上の様々な問題の最適化に関連したアルゴリズムに関する研究を、数理学・理論計算幾何学・情報工学・社会科学など側面から行っております。これらの分野の知見を統合し、様々な応用に対するアルゴリズム成果の適用を行うとともに、コアアルゴリズムの先進的なライブラリの構築を目指しています。

また、ネットワークアルゴリズムに関する境界の数理的証明と、そこでの開発技術の情報システムへの利用、特にインターネット上でのデータ検索技術への応用なども行い、数理学から実世界ネットワークにわたる研究を総括する予定です。

グループリーダーは徳山豪です。他グループに比べてこれまで活動頻度が少なかったのですが、メンバーは徐々に増えており、これから活発に活動していく予定です。現在のメンバーは2013年2月にスイスから来たLorenz Klausに加え、昨年12月には元徳山研メンバーのMatias Kormanがスペイン(カタルーニャ)から復帰し、今年9月からはオランダ人のAndreas van Renssenがカナダよりやってきました。このような国際色豊かなネットワークアルゴリズムグループです(余談ですが、徳山研究室には現時点で9カ国の教員・学生がいて、実に国際的です)。

現時点での当グループの特徴は、計算幾何学*のスペシャリストが多いことです。この利点を生かし、アドホックネットワーク等の各種ネットワークに対し、計算幾何学的アプローチによるアルゴリズム開発を目指しています。(塩浦昭義)

計算幾何学 Computational Geometry

幾何学の言葉で述べることのできるアルゴリズムの研究をテーマとする計算機科学の一分野。

検索技術 Search Technology

データの集合の中から目的とするデータを探し出すこと。コンピュータやインターネットの発達により、検索技術の重要性は近年増している。



グループミーティングの様子(東北大学)



サブリーダー

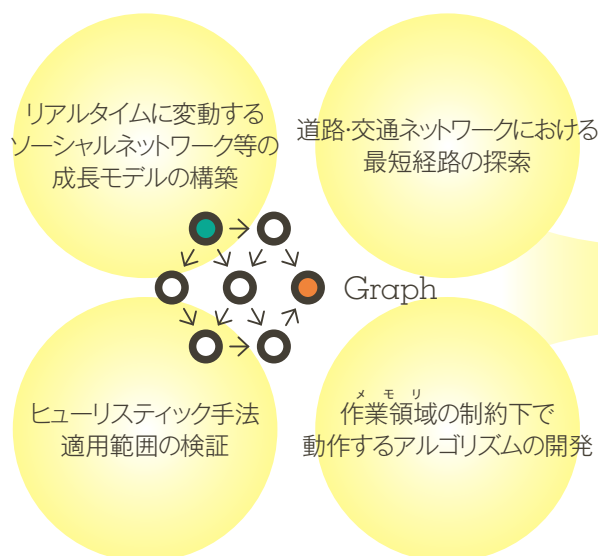
塩浦昭義 Akiyoshi SHIOURA

東北大学大学院情報科学研究科 准教授/国立情報学研究所 客員准教授

「巨大グラフ」を解析可能な革新的・高速アルゴリズム開発

Developing high-speed data processing algorithms.

「理論計算機科学」「離散数学」を基礎とした……



新たな数学的理論を構築、
ネットワーク解析における
理論的研究の有効性を実証

巨大情報量の解析を伴う
社会の諸問題解決の糸口に

インターネットのウェブ構造や、Facebook、Twitterなどのソーシャルネットワークに代表される巨大なネットワークは、各々 10^9 (10億)人に近いユーザーが利用し、現代社会に欠かせない存在となっています。これらのネットワークは年々急速に膨張し、近い将来には 10^{10} (100億)を超えるサイズになると予想されています。

ネットワークの膨張に伴う情報量の増大はハードウェアの進歩を上回る速さで進んでおり、いわゆる「ビッグデータ」の中でも特に巨大な、 10^{10} 以上のサイズのネットワークに対しては、現行のアルゴリズムでは実用的な速度で情報を解析することが不可能であり、高速アルゴリズムの開発が急務となっています。このような背景のもと、本プロジェクトでは、巨大なネットワークを膨大な点と辺の接続構造、すなわち 10^{10} 以上の頂点を持つ「巨大グラフ」として表現し、理論計算機科学や離散数学などにおける最先端の数学的理論を駆使してそれを解析する、高速アルゴリズムの開発を目指します。

研究成果

論文名	著者	掲載誌
<i>A Simple Algorithm for the Graph Minor Decomposition - Logic Meets Structural Graph Theory</i>	Martin Grohe Ken-ichi Kawarabayashi Bruce Reed	Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA2013), 2013, 414-431
<i>All-or-nothing multicommodity flow problem with bounded fractionality in planar graphs</i>	Ken-ichi Kawarabayashi Yusuke Kobayashi	Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2013), 2013, 187-196
<i>Dynamic and Historical Shortest-Path Distance Queries on Large Evolving Networks by Pruned Landmark Labeling</i>	Takuya Akiba, Yoichi Iwata Yuichi Yoshida	Proceedings of the 23rd International Conference on World Wide Web (WWW 2014), 237-248
<i>Almost Linear-Time Algorithms for Adaptive Betweenness Centrality using Hypergraph Sketches</i>	Yuichi Yoshida	Proceedings of the 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2014, 1416-1425
<i>4-connected projective-planar graphs are hamiltonian-connected</i>	Ken-ichi Kawarabayashi Kenta Ozeki	Journal of Combinatorial Theory Series B *to appear
<i>The Parameterized Complexity of k-Biclique</i>	Bingkai Lin	Proceeding of the 26th ACM-SIAM Symposium on Discrete Algorithms (SODA2015) *to appear

主要受賞

受賞年月	受賞名	受賞者
2013年 1 月	Best Paper Award ACM-SIAM Symposium on Discrete Algorithms (SODA2013) <i>A Simple Algorithm for the Graph Minor Decomposition - Logic meets Structural Graph Theory</i>	Martin Grohe Ken-ichi Kawarabayashi Bruce Reed
2015年 1 月	Best Paper & Best Student Paper Award ACM-SIAM Symposium on Discrete Algorithms (SODA2015) <i>The Parameterized Complexity of k-Biclique</i>	Bingkai Lin
2013年 2 月	第9回(平成24年度)日本学士院学術奨励賞&日本学術振興会賞	河原林健一
2013年 12月	井上科学振興財団 第30回井上研究奨励賞奨励賞	吉田悠一
2013年 9 月	2013年度日本数学会賞 建部賢弘賞 奨励賞	小関健太
2014年 8 月	日本オペレーションズ・リサーチ学会 2014年度研究賞奨励賞	福永拓郎



キックオフ
(東京・千代田区/2013年3月)



中間発表会 2013年夏
(栃木・那須/2013年8月)



成果報告会 2013
(東京・千代田区/2014年3月)



感謝祭 Summer 2014
(東京・千代田区/2014年8月)



中間発表会 2014
(静岡・伊東/2014年9月)



[オーガナイズ] 日仏ビッグデータワークショップ
NII / NICT / フランス大使館共催(東京・港区/2014年11月)



[オーガナイズ] 日英ビッグデータ研究ワークショップ
NII / 英国大使館共催(東京・港区/2014年2月)



[基調講演]
Japan Science Day
(ニューデリー、インド/2014年1月)



[基調講演] 「グローバルな研究が生む知の潮流」
都立戸山高専・生徒研究成果合同発表会
(東京・新宿区/2014年2月)



[講演]
統数研・数学協働プログラム
チュートリアル
ビッググラフと最適化
(東京・立川/2014年3月)



プロジェクト紹介
ムービー制作中

本気の
プログラミング
コンテスト対決

アルゴリズムって
何?

初心者にも
やさしい
解説

2015年2月、
ホームページで
公開予定

共同研究パートナー企業(五十音順)

NEC / NTTコミュニケーション科学基礎研究所 機械学習・データ科学センター / 株式会社地球最適化インスティテュート(三菱ケミカルホールディングス) / 株式会社Preferred Networks / 日本アイ・ビー・エム株式会社 東京基礎研究所 / ヤフー株式会社 Yahoo! JAPAN 研究所



ERATO 河原林巨大グラフプロジェクト事務局

〒101-8430 東京都千代田区一ツ橋2-1-2 学術総合センター内 国立情報学研究所 ビッグデータ数理国際研究センター

☎03-4212-2664 ✉kk-erato-jimu@nii.ac.jp 🌐http://bigdata.nii.ac.jp/ 📘www.facebook.com/erato.kawarabayashi