

**離散数学、アルゴリズムの新研究室！**

**研究の進め方等は、個人に合わせていき  
たいと思います。**

**最適な研究環境を提供します！**

**経済的なサポートもなるべくする予定！**

# 河原林健一

◆ 専門分野(詳細): グラフ理論(離散数学), アルゴリズム  
理論計算機科学、グラフアルゴリズム, 高速アルゴリズム開発

◆ 国際会議: STOC, FOCS, SODA(理論), LICS(論理)  
ICML, NeuRIPS, ICLR, AISTAT(機械学習),  
SIGMOD, VLDB, ICDE(データベース),  
KDD, WWW, CIKM(データマイニング), AAAI, IJCAI(AI), ACL(自然言語)

◆ 研究歴:  
~2005年: グラフ理論、離散数学 → 数学のみ  
2005~2011年: グラフ理論, アルゴリズム → 数学&CS  
2011年~ グラフ関連何でも!  
2012~2019年: JST ERATO「河原林巨大グラフ」PD 16億  
2018~2026 : 基盤S PI 2億(5年)  
2019~ : JST ACT-X「数理と情報」研究総括

◆ 受賞歴  
2008年: 日本IBM科学賞  
2013年: 日本学術振興会賞・日本学士院奨励賞、SODA'13 BEST PAPER  
2015年: 日本数学会春季賞  
2021年: Fulkerson Prize(離散数学、アルゴリズムで世界で最も権威ある賞)



NII教授

所在地: 神保町・竹橋(NII)  
本郷

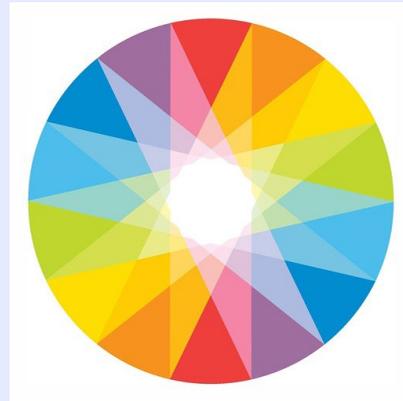
NII 副所長

2019-2021

Shonan Meeting Chair

2017-

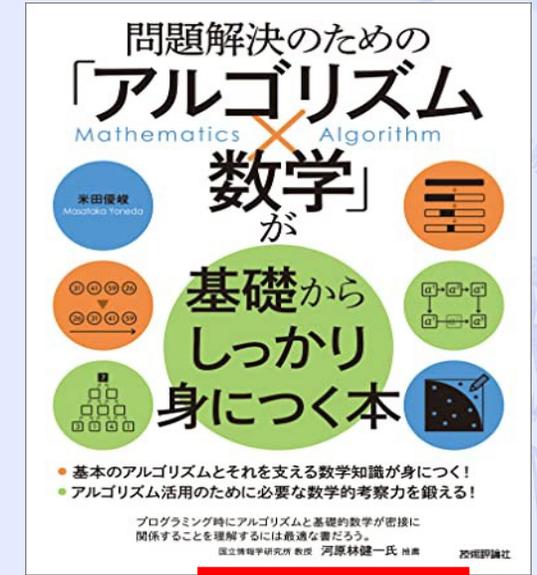
# アルゴリズムと離散数学研究室 (BRAND NEW!)



情報科学の達人主催



この本を買わずして何をかう！とは言ってない。  
小さいところだけ



教授: 河原林 健一 (多分左)  
NII (国立情報学研究所) 教授  
所在地: 神保町・竹橋 (NII)



NII

## ◆ 研究トピック (アルゴリズム全般)

1. グラフ理論、離散数学
2. グラフアルゴリズム → **理論的に速いアルゴリズム開発**  
グラフカット、グラフ彩色、フローなど
3. 高速アルゴリズム → **実装して速いアルゴリズム**  
最短パス、検索など
4. グラフ関連何でも → **深層学習 (GNN), 自然言語処理**  
**データ解析などへの応用**
5. データ構造・組合せ最適化・理論計算機科学

**離散数学、アルゴリズムの新研究室！  
でもなぜアルゴリズム、離散数学を研究  
するの？**

---

# アルゴリズム&離散数学研究

## 学術的意義

- ◆ 数学・計算機科学分野の超難問とブレイクスルーの要請
  - P vs NP (7大数学難問の一つ)
  - Shorの素因数分解の多項式「量子的」アルゴリズム  
→ 現在の「量子コンピューター」ブームのきっかけ！
- ◆ ハードウェアの限界 アルゴリズムからのスピードアップ！
  - ムーア則の限界から、ハード面(特にCPU)で100倍スピードアップ達成は難しい。
  - しかしアルゴリズムでは100倍スピードアップは日常茶飯事

# アルゴリズム&離散数学研究

## 社会的意義

- ◆ アルゴリズム理論の導入による新しい実用手法がGAFAを作り、IT社会を先導！
  - 検索エンジン (HITS, PAGERANKなど)
  - 暗号 (RSA暗号など)
  - プライバシー (差分プライバシーなど)
  - バイオインフォマティクス
- ◆ 機械(深層)学習ブーム
  - Explainable AI (アルゴリズムの解析が必要！)
  - スケール化とアルゴリズムの効率化が必要！
- ◆ 量子コンピューターブームのきっかけ
  - Shorの素因数分解の多項式「量子的」アルゴリズム

# 離散数学、アルゴリズム 理論研究の例

---

# アルゴリズム・離散数学研究の例

**Min-Cut:** グラフアルゴリズムで最も歴史があり、有名な問題

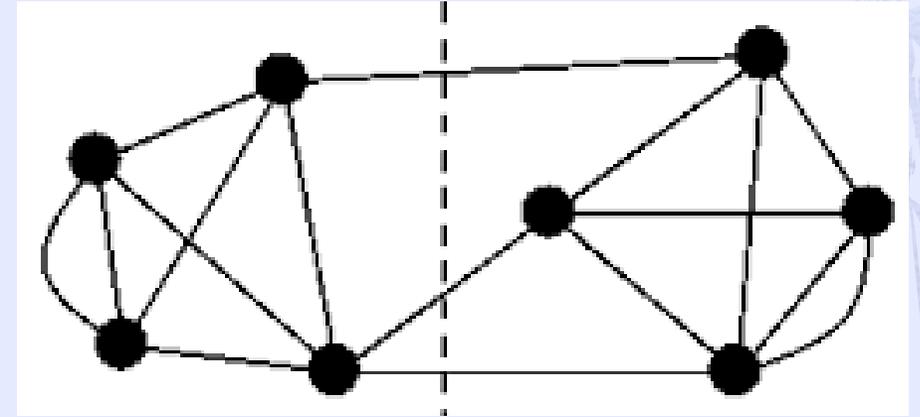
**具体的貢献:** 決定的(deterministic)な、  
ほぼ線形アルゴリズムを開発！

(Journal of ACM 2018)

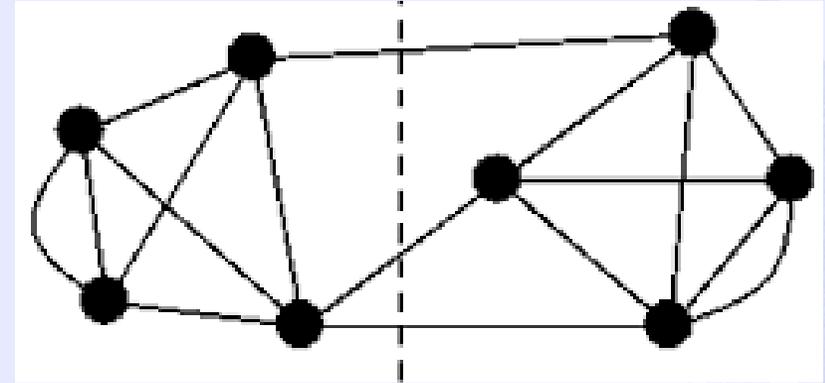
→ Fulkerson Prize 2021!

アルゴリズム、組合せ最適化、ネットワーク、グラフ理論などにおける  
当該分野のトップ国際賞

Min-cutの歴史は？



# Min-cutの歴史



$m, n$ : # of vertices, edges, respectively.

**Race** by many “giant” researchers in theory !

Ford and Fulkerson '56  $O(mn^2)$  ← **ORの創立者!**

Even and Tarjan '75  $O(nm^{1.5})$  ← **Turing賞受賞者**

Karzanov and Timofeev '86  $O(n^3)$  ← **ロシア学派**

Hao and Orlin '92  $O(nm \log(n^2/m))$  **Deterministic**

Nagamochi and Ibaraki '92  $O(nm+n^2 \log n)$  **Deterministic**

Karger '94  $O(mn^{0.5})$  **(Randomize, Monte Carlo) J.ACM**

Karger and Stein '96  $O(n^2 \log^3 m)$  **(Randomize, Monte Carlo)**

Karger '99  $O(m \log^3 n)$  **(Randomize, Monte Carlo) J. ACM**

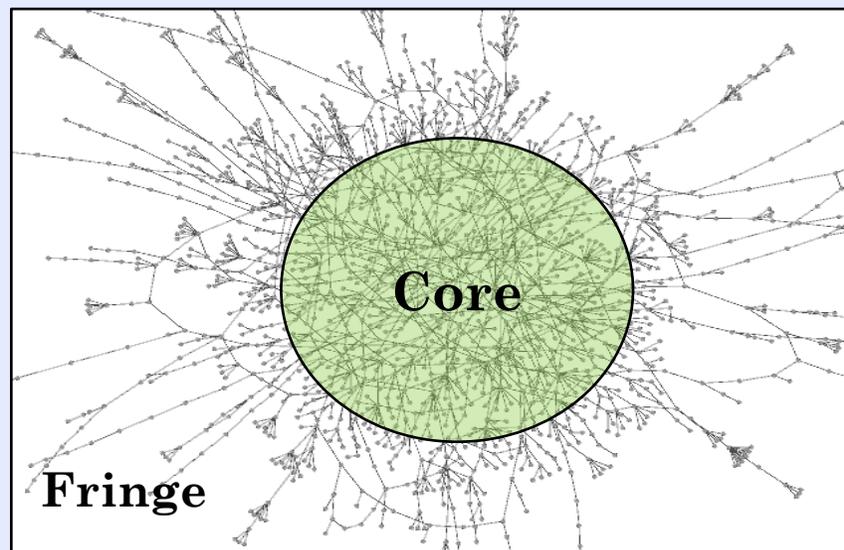
**KK and Thorup (STOC '15)  $O(m \log^5 n)$  Deterministic! J.ACM**

**実世界の巨大グラフにおける  
高速アルゴリズム開発：  
離散数学・アルゴリズムの応用**

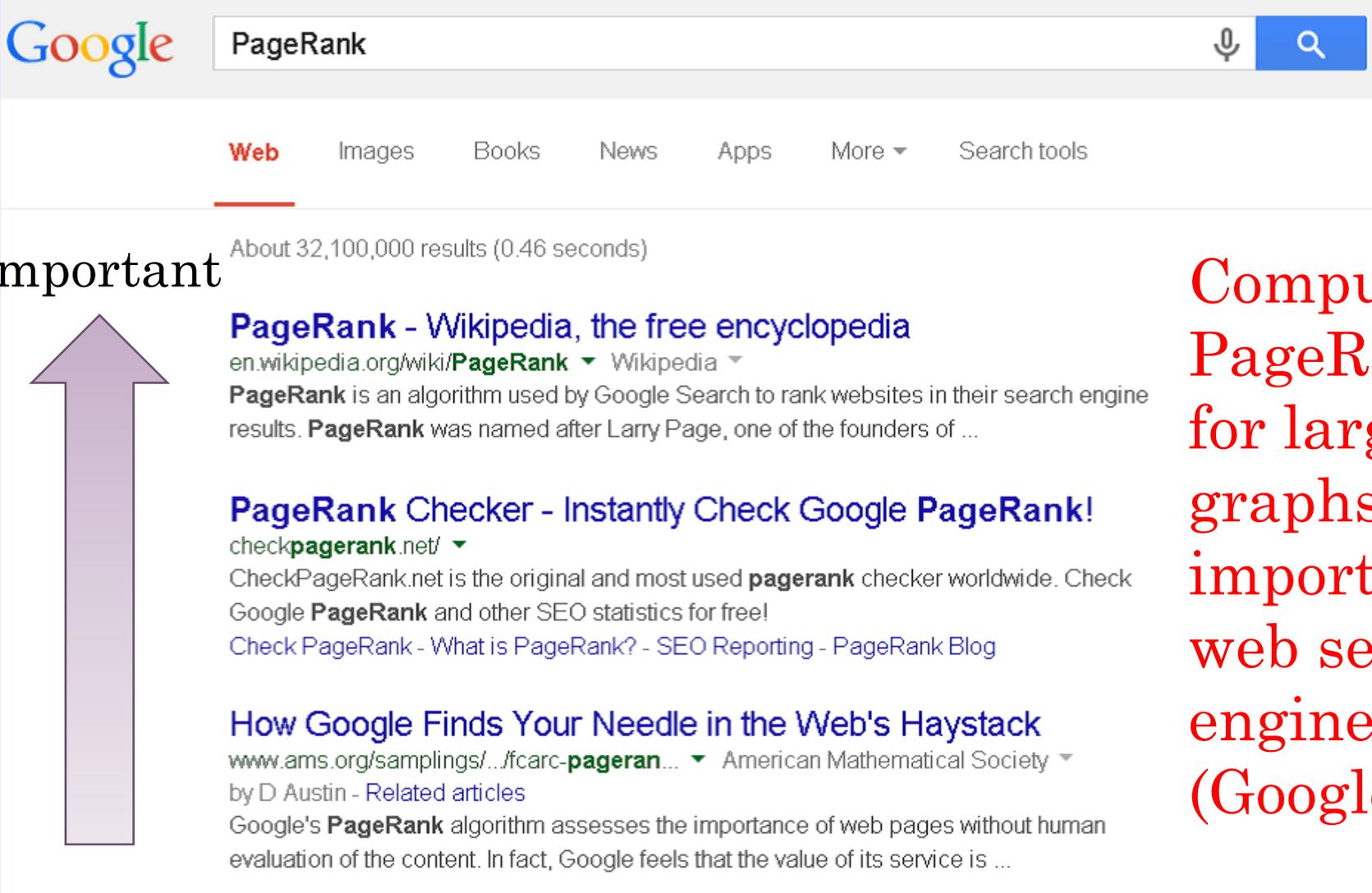
---

# 基礎研究の貢献：ネットワークの構造利用

- ◆ 現実のソーシャルネットワーク
  - 低次数の頂点が縁でツリー状(木分解)になっている構造
- ◆ この性質を積極的に利用
  1. PageRankの計算(通常3-5倍高速化) (VLDB'14)
  2. SIMRANKの掲載(700-1000倍)



# Vertex Ranking Measure



The image shows a screenshot of a Google search for "PageRank". The search bar at the top contains the text "PageRank" and a search icon. Below the search bar, there are navigation tabs for "Web", "Images", "Books", "News", "Apps", "More", and "Search tools". The search results are displayed below, showing "About 32,100,000 results (0.46 seconds)". The first result is "PageRank - Wikipedia, the free encyclopedia" with a snippet: "PageRank is an algorithm used by Google Search to rank websites in their search engine results. PageRank was named after Larry Page, one of the founders of ...". The second result is "PageRank Checker - Instantly Check Google PageRank!" with a snippet: "CheckPageRank.net is the original and most used pagerank checker worldwide. Check Google PageRank and other SEO statistics for free!". The third result is "How Google Finds Your Needle in the Web's Haystack" with a snippet: "Google's PageRank algorithm assesses the importance of web pages without human evaluation of the content. In fact, Google feels that the value of its service is ...". A large purple arrow points upwards from the bottom left towards the word "Important" and the first search result. The word "Important" is written in a large, black, serif font. The entire screenshot is overlaid on a light blue background with faint floral patterns.

Important

About 32,100,000 results (0.46 seconds)

**PageRank - Wikipedia, the free encyclopedia**  
en.wikipedia.org/wiki/**PageRank** ▾ Wikipedia ▾  
**PageRank** is an algorithm used by Google Search to rank websites in their search engine results. **PageRank** was named after Larry Page, one of the founders of ...

**PageRank Checker - Instantly Check Google PageRank!**  
check**pagerank**.net/ ▾  
CheckPageRank.net is the original and most used **pagerank** checker worldwide. Check Google **PageRank** and other SEO statistics for free!  
Check PageRank - What is PageRank? - SEO Reporting - PageRank Blog

**How Google Finds Your Needle in the Web's Haystack**  
www.ams.org/samplings/.../fcarc-**pageran**... ▾ American Mathematical Society ▾  
by D Austin - Related articles  
Google's **PageRank** algorithm assesses the importance of web pages without human evaluation of the content. In fact, Google feels that the value of its service is ...

Computing PageRank for large web graphs is very important for web search engine!  
(Google etc)

# 線形連立方程式 $Ax=b$

代表的応用: PageRank(重要度)

$$x = (1-a) Ax + a b$$

- ◆  $A$ は隣接行列！
- ◆ 各ユーザは確率  $a$  でホームページ  $b$  に戻り, 確率  $(1-a)$  で次のページに進む
- ◆  $x$  の値 = そのノードにいる確率 = 重要度

$A$  のグラフ構造を利用して効率的に  
PageRank (逆行列) を計算する！

# 木分解を利用したPageRank 計算

28/43

従来手法:

- ◆ 反復法:「 $x \leftarrow (1-a) A x + a b$  を繰り返す」
  - ・早くはないが, 大規模なグラフでも動く
- ◆ 直接法:「ガウスの消去法で解く」
  - ・基本的に動的計画法
  - ・実際のウェブでは, 計算が終わらない



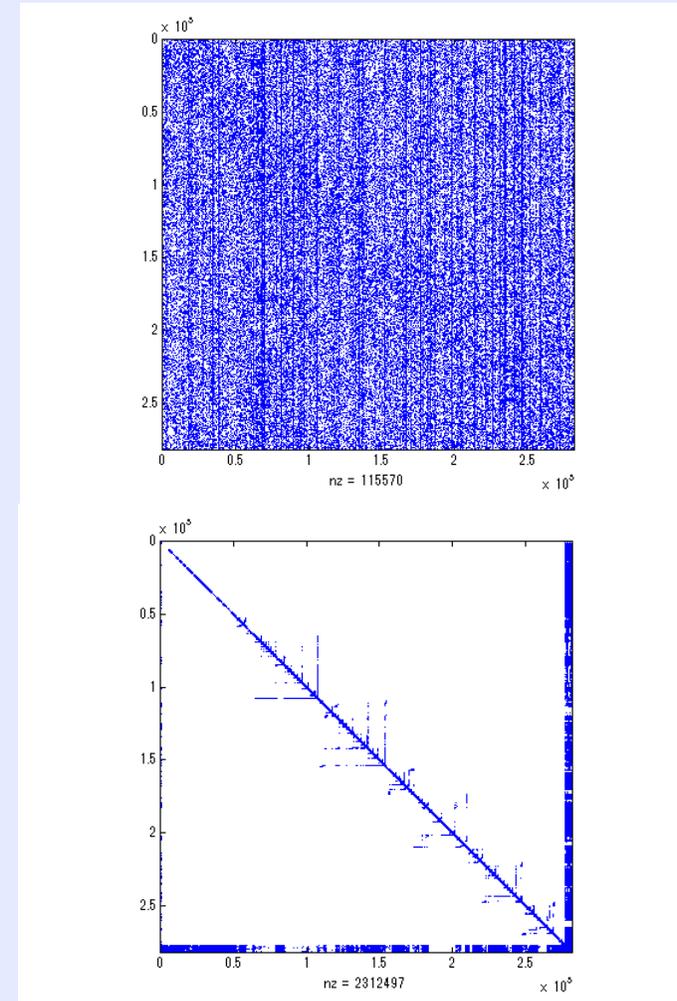
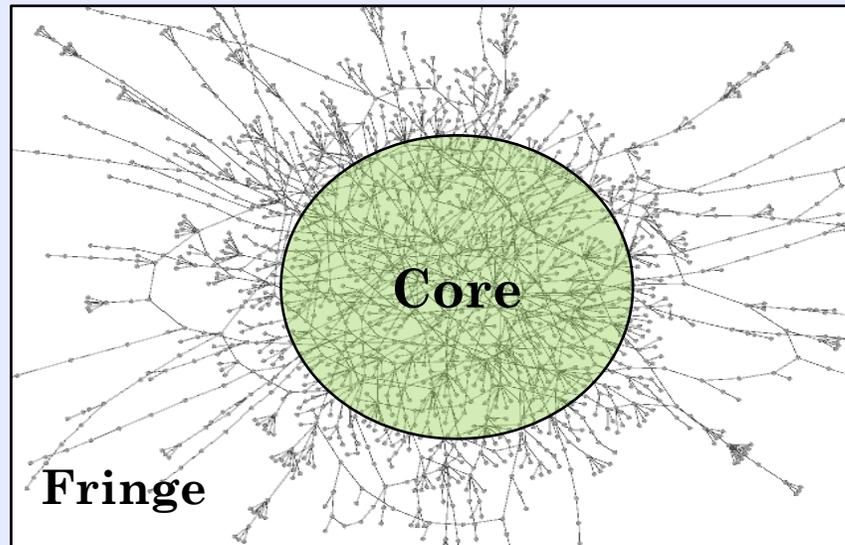
提案手法:

グラフの木分解を「前処理」で行い,  
コア以外は直接法, コアは反復法で解く!

# ウェブグラフの木分解

29/43

- ◆ コア以外: 「木幅が小さい」
  - 直接法で早い!
  - 動的計画法で動く!
- ◆ コア: 「エクспанダー」
  - 直接法は遅い!
  - 反復法がすばやく収束

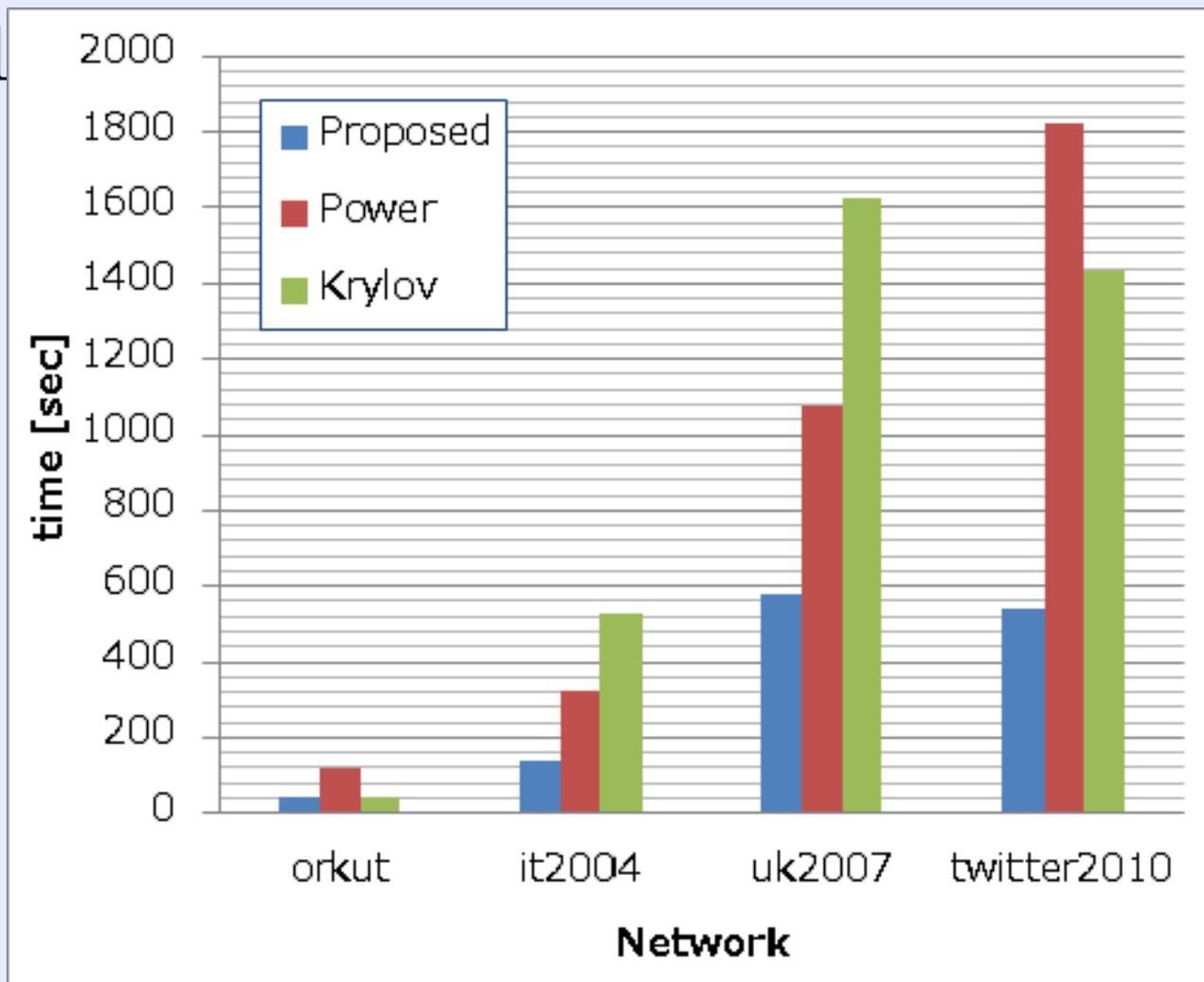


web-Stanford (281k nodes, 2M edges)

# Our Contribution

- ◆ A fast and accurate algorithm for real-world networks

◆ 1



twitter-2010

$64 \times 10^6$  vertices

$1 \times 10^9$  edges

uk2007

$105 \times 10^6$  vertices

$4 \times 10^9$  edges

**実世界のデータに対する  
高速アルゴリズム開発：  
数理学の応用例**

---

# 研究成果(例) Twitterのリアルタイム解析

Twitterをリアルタイム解析  
情報学研

## 話題抽出、毎分6万件

### スパム・顔文字自動削除

国立情報学研究所の林浩平特任助教、河原林健一教授らの研究チームは、ツイッタをリアルタイムに解析し、活発な話題を抽出する技術を開発した。毎分約6万件のペースで発信される日本全国のツイートを即時に分析できる。今、この瞬間に世界で起きている出来事を直ちに把握できるようになる。

林特任助教らは、モラルゴリズムと呼ぶるたびに、繰り返し膨大な分析が行える。ノットを関連づける計算手法を開発した。大なる計算を必要とするツイッタのユーザ「関係データ解析」のある一定の期間にあり、リアルタイムな新しい手法として、途切れることなく、同時に処理が困難だった。また、世界中に数億人いるユーザーのツイートをリアルタイムに解析できる。ストリークに出現するキーワードを一括送信される「スパム」を排除することも、この瞬間に世界で起きている出来事を直ちに把握できるようになる。

トピック抽出の可視化例。文字が大きいほど頻出度が高い(情報学研提供)

日刊工業新聞  
平成27年8月11日(火)朝刊1面

- ◆ データマイニングトップ会議のKDDで発表  
**Hayashi, K., Maehara, T.,**

**Toyoda, M. & Kawarabayashi, K.**

Real-time Top-R Topic Detection on Twitter with Topic Hijack Filtering  
ACM SIGKDD Conference on Knowledge Discovery and Data Mining  
(KDD), 2015

- ◆ その後、日刊工業新聞の一面に掲載

# 取り除かれた不正トピックの例

- シネマトゥデイ 妖精トム 映画 ハリウッド 戦闘 クルーズ 主演
- 発売 2013 情報 開催 参加 商品 入荷
- 限定 情報 開催 イベント 好評 商品 @\*\*\*\*\*
- リプライ 富士山 樋口 フェスティバル 興味 早稲田 開催 現在 不問 募集 スタッフ 大

広告系

- リフォロー アカウント 19836 フォロワー Only 希望 支援 696382 交流
- 無料 日本 拡散 希望 2013/04 リブ

拡散系

- 特徴 燃料 格闘 操縦 射撃 装甲 整備 評価 機動
- 所有 City ポイント 前回 Tweet アクセス 獲得 Intel
- 自動 だれ 宣伝 AutoTweet オートツイート Twitter 定期 設定 サイト

ボット系

- とノ ω・ ㄣ | ㄣㄣ | ㄣ ㄣ | ㄣ ㄣ | ㄣ ㄣ | ㄣ
- ㄣ | う ㄣ ㄣ Å・ | お ● ㄣ ㄣ | よ。) ㄣ。・
- \*・ .\* :.。... :. \*! \*!° :.。... :.

顔文字系

# 著者の感想

「理論的にはあまりたいしたことない。  
でもうまく「数理科学」の知識を  
組みあわせることができた。」

→数理科学の知見を効果的に  
組み合わせることで可能に！  
学生さんの「推進力(実装力)」も大きな武器に！  
産業界から問い合わせ多数！

# その他の具体的な研究対象

---

## 1. グラフ理論、離散数学

4色定理とその拡張、グラフ構造に関する研究など

## 2. グラフアルゴリズム → 理論的に速いアルゴリズム開発

グラフ彩色問題 (グラフ彩色に関する近似アルゴリズム開発) など

## 3. グラフ関連何でも

深層学習 (GNN), 自然言語処理、データ解析、実ネットワークへの応用。グラフ理論、グラフアルゴリズムなどの応用。

## 4. データ構造・組合せ最適化・理論計算機科学

**離散数学、アルゴリズムを学び、  
応用したい学生さんへ！**

---

# 離散数学&アルゴリズムを学ぶことにより、 (実用)研究に参加するメリット

1. 「最適化」問題をセンスよく、かつ効率的に解く技術を多数もっている。
2. データサイエンスから派生する問題を、(近似的に)解決可能に「定式化」する技術を多数持っている。特に物理学者などが作った「モデル」の精神を引き継ぎながら、計算可能なモデルに変更する能力に長けている。  
→ アルゴリズムを動かすときに、式の魂の叫び&パラメーターの悲鳴が聞こえる！行列&グラフ&多面体の形が見える！
3. 幅広い数学の道具を、「広く」&「浅く」組み合わせることができる！  
→ 離散数学、アルゴリズムの専門家のみが可能！これこそ汎用性！

# 意欲的な皆さん、一緒に研究しましょう！

1. 「難しい」研究課題に対して
2. 「理論ベース」アプローチで
3. 「未来」の人材(5年後に世界的に競争力ある世代)が
4. 「国際的な」研究成果を見せる！

そのために必要なこと

1. 離散数学(グラフ理論、組合せ論)、グラフアルゴリズム、組合せ最適化等の最先端の成果と知見を得る←私の専門！
2. 優秀な大学院生を巻き込む  
→ スーパーエリートへの道筋を作る！

日本のVISIBILITY向上へ！

END

Question ?

