

令和 5 年 6 月 2 日現在

機関番号：34310

研究種目：若手研究(B)

研究期間：2017～2022

課題番号：17K12794

研究課題名（和文）スパース最適化に基づく時系列ネットワークからの研究トレンドマッピング

研究課題名（英文）Mapping emerging research trends from dynamic networks via sparse representation

研究代表者

桂井 麻里衣 (Katsurai, Marie)

同志社大学・理工学部・准教授

研究者番号：70744952

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究課題の目的は、学術ビッグデータからの最新トピック発掘とその変遷の可視化を同時実現する研究トレンドマッピング技術の確立であった。研究期間では、時系列ネットワークにおいて急激に時間変化した部分のみを残すスパース最適化手法を提案した。得られた成果は学術データ分析に関する英文論文誌Scientometricsで発表し、研究代表者のウェブサイトソースコードを公開した。また、インタフェースへの応用成果は国際会議JCDL2021でデモ論文として発表した。

研究成果の学術的意義や社会的意義

過去の状態に比べて急激に変化したノードとエッジを可視化するという目的におけるスパース表現の採用は本研究の独創的な点であり学術的な意義が大きい。各研究分野で急成長している技術やその発展の様子を解明することは、研究関係者や政策立案者の俯瞰的視野の養成につながる。本研究で構築した手法をウェブ上で公開することで、関連分野の技術開発促進に貢献した。

研究成果の概要（英文）：The objective of this research project was to establish a research trend mapping technique that enables simultaneous exploration of the latest topics and visualization of their evolution from academic big data. During the research period, we proposed a sparse optimization method that preserves only the rapidly changing parts in the time-series of co-word networks. The obtained results were published in Scientometrics, and the source code was made publicly available on the researcher's website. Furthermore, the application results were presented as a demo paper at JCDL 2021.

研究分野：マルチメディア情報検索

キーワード：学術データ分析 サイエンスマッピング トレンド分析 トレンド可視化

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

国際会議やジャーナル論文の継続的増大、プレプリントサーバの普及は、学術情報のビッグデータ化をもたらした。急速に発展する科学技術の動向を把握し、研究企画戦略を適切に行うには、これらのビッグデータから研究分野の知識構造を抽出・可視化する技術の確立が急務である。これまで、用語、論文、研究者などをノード(頂点)とし、それらの関係をエッジ(辺)で表す複雑ネットワークの有効性が示されてきた。代表的な複雑ネットワークとして、図1に示すように、用語の共起関係や研究者の共著関係、論文の引用関係などが挙げられる。各ネットワークはそれぞれ異なる知識構造を表出する。例えば用語の共起関係は研究課題や科学技術の関連性を示し、論文の引用関係は研究領域の関係を表す。また、共著関係からは中心研究者や主要な研究グループを発見できる。

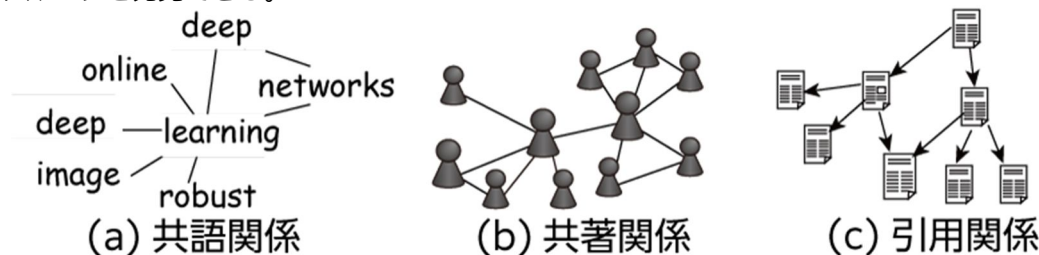


図1 学術データから構築されるネットワークの例

上記の書誌・引用情報に基づく関係を超えて、研究者の間の潜在的な類似性を発見することも、共同研究の活性化支援となる。申請者はこれまでに、国内の全研究分野を包括した学術論文データベースを用いて研究者の専門トピックを算出する手法を構築した。また、研究者間の専門トピックの類似関係をネットワークで表現すると、過去に共著経験がなくとも共同研究に適した関連研究者を推薦できることを明らかにした。以上の成果を発展させ、研究者らが注目する最新のトピックを発掘するためには、ネットワークに時間軸の概念を導入すべきとの着想に至った。

2. 研究の目的

本研究の目的は、学術データから構築されるネットワークを情報源としたトレンドトピック抽出手法の構築である。任意の複雑ネットワークを時間区間ごとに構築すると、図2に示すような時系列ネットワークが得られ、これを情報源として研究トレンドを抽出する試みが報告されている。既存の手法は、図2の点線枠内に示すように、まずネットワークの構造に基づき各ノードの重要度を算出し、その時間変化を追跡する。そして、過去に比べて重要度が上昇したノードを、その時間区間におけるトレンドトピックとみなす。しかしながら、この方法ではノードの重要度という数値のみを追跡するため、出力されるトピック間の関連性は不透明となる。したがって、従来アプローチでは、研究トレンドに該当する知識構造を適切に抽出できていないといえる。

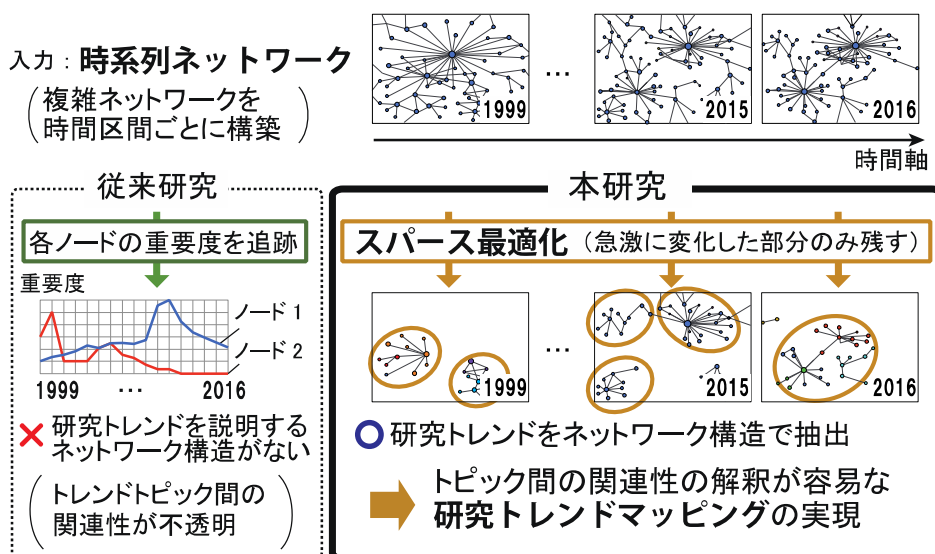


図2 本研究の目的

従来アプローチの問題を解決するには、研究トレンド自体をネットワーク構造で表現するのが有効と考えられる。そこで本研究課題では、図1の太枠内に示すように、もとの複雑ネットワークを間引いて研究トレンドに該当する部分のみを残すスパース表現に着目する。ネットワークのスパース表現を得るためには、ある基準(目的関数と呼ばれる)の下、不要なノードとエッジを決定するスパース最適化が必要となる。過去に比べて顕著に重みが増加したノードとエッジを残すように目的関数を設計することで、各時間区間に特有のネットワーク構造が抽出できると考えられる。このようなスパース最適化手法を構築することで、トピック間の関連性の解釈が容易な研究トレンドマッピングの実現を目指す。

3. 研究の方法

手法構築、評価実験、応用という三段階で研究を進めた。手法構築は目的関数の設計から始めた。任意の時系列ネットワークから、行をエッジの重み、列を時間区間とした行列を算出した。得られた行列を緩やかな時間変化をもつ行列と急激な時間変化をもつ行列に分解した場合の誤差関数を目的関数とした。近接勾配法と呼ばれる反復法の高速版である高速繰り返し縮小法を適用することで、高速かつ省メモリな手法を実現した。

本研究における研究トレンドマッピングは、正解データを用意することが困難であった。よって、手法の定量評価のために、人工データを用いた数値シミュレーションを実施した。具体的な評価方法として、既知の確率分布を用いて作成した時系列ネットワークにおいて、人為的に急激に変化させた部分をトレンドとみなして検出精度を算出した。また、人工データにノイズを混ぜて実験を行い、手法の頑健性を評価した。評価指標によって提案手法の有効性を確認したあとは、ウェブ上の学術論文データベースから取得した実データに手法を適用し、定性的評価を行った。

最後に、本研究成果による情報検索への応用として、二つの論文誌間のトレンドを提示し、それらの間の共通点を可視化するインタフェースを構築した。

4. 研究成果

ベースライン手法に比べて高速かつ高精度にバースト検出可能な手法を構築できた。定性評価は主に図1(a)に示した語の共起関係を対象に行った。提案手法による得られたトレンド共起語ネットワークは、二語以上からなる技術名や技術間の関連性を概念空間で直接表現できるという利点をもつ。

スパース最適化は動画圧縮や脳活動のモデリングなど様々な問題において効果が実証されてきたが、研究トレンド抽出に用いた例は存在しなかった。特に、過去の状態に比べて急激に変化したノードとエッジを可視化した点は、既存のスパース最適化応用には見られず、本研究の学術的な特色といえる。

研究成果は国内外の学会で発表し、そこでのフィードバックを受けて取りまとめた論文は学術データ分析に関するトップ論文誌の一つである *Scientometrics* でオープンアクセス形態にて発表した。異なる国際会議の共通トレンド可視化インタフェースへの応用は、デジタルライブラリの有名国際会議の一つである JCDL2021 にてデモ論文として発表した。作成したプログラムやトレンド可視化インタフェースのデモ動画を研究代表者のウェブサイトで公開するなど、オープンサイエンスに貢献した。

本研究を通じて得たトレンド検出の知見をもとに、様々な分野のキーワードのバースト性を定量化した。その一環として、図書館情報学分野で盛んに用いられているデータ分析手法の解明に取り組み、その成果は国際共著論文として公表した。

上記の研究トレンドの分析・可視化の成果を得ただけでなく、学術データの分析と学習に関する最新動向の理解や新たな問題意識の発見につながった。それらの一部は ACM SIGWEB Newsletter の招待論文にて概説した。

以上の学術的な発信のみならず、オープンキャンパスでの研究発表やデモ展示を通じ、中高生などの若い世代に研究成果を示した。今後も同様のイベントにて積極的に成果を公開することで、非専門家に対する学術研究のオープン性を高め、当該技術分野の重要性を広くアピールする予定である。

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 3件/うち国際共著 1件/うちオープンアクセス 4件）

1. 著者名 Marie Katsurai, Shunsuke Ono	4. 巻 121
2. 論文標題 TrendNets: mapping emerging research trends from dynamic co-word networks via sparse representation	5. 発行年 2019年
3. 雑誌名 Scientometrics	6. 最初と最後の頁 1583--1598
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s11192-019-03241-6	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Marie Katsurai	4. 巻 4
2. 論文標題 Using word embeddings for library and information science research: a short survey	5. 発行年 2020年
3. 雑誌名 ACM SIGWEB Newsletter	6. 最初と最後の頁 Article No. 4
掲載論文のDOI（デジタルオブジェクト識別子） 10.1145/3387726.3387730	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 荒木将貴, 桂井麻里衣, 大向一輝, 武田英明	4. 巻 16-J
2. 論文標題 大学における部局横断型共同研究の活発さを把握する指標の検討	5. 発行年 2018年
3. 雑誌名 日本データベース学会和文論文誌	6. 最初と最後の頁 1-6
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Marie Katsurai and Soohyung Joo	4. 巻 19
2. 論文標題 Adoption of Data Mining Methods in the Discipline of Library and Information Science	5. 発行年 2021年
3. 雑誌名 Journal of Library and Information Studies	6. 最初と最後の頁 1-17
掲載論文のDOI（デジタルオブジェクト識別子） 10.6182/jlis.202106_19(1).001	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

〔学会発表〕 計7件（うち招待講演 1件 / うち国際学会 5件）

1. 発表者名 Tomoya Nishide and Marie Katsurai
2. 発表標題 Finding the Relevance Between Publication Venues Based on Research Trend Similarity and Citation Relationships
3. 学会等名 2021 ACM/IEEE Joint Conference on Digital Libraries (JCDL) (国際学会)
4. 発表年 2021年

1. 発表者名 丹後綱也, 西澤浩之, 近澤悠登, 桂井麻里衣
2. 発表標題 関連研究者と所属位置情報の検索・可視化システム
3. 学会等名 第12回データ工学と情報マネジメントに関するフォーラム (DEIM2020)
4. 発表年 2020年

1. 発表者名 Mayu Kimura, Marie Katsurai
2. 発表標題 Investigating the Consistency of Emoji Sentiment Lexicons Constructed Using Different Languages
3. 学会等名 The 20th International Conference on Information Integration and Web-based Applications & Services (iiWAS2018) (国際学会)
4. 発表年 2018年

1. 発表者名 Hiroyuki Nishizawa, Marie Katsurai, Ikki Ohmukai, Hideaki Takeda
2. 発表標題 Measuring Researcher Relatedness with Changes in Their Research Interests
3. 学会等名 018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) (国際学会)
4. 発表年 2018年

1. 発表者名 Mayu Kimura and Marie Katsurai
2. 発表標題 Automatic Construction of an Emoji Sentiment Lexicon
3. 学会等名 Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 Satoshi Sanjo and Marie Katsurai
2. 発表標題 Recipe Popularity Prediction with Deep Visual-Semantic Fusion
3. 学会等名 Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (国際学会)
4. 発表年 2017年

1. 発表者名 桂井麻里衣
2. 発表標題 異分野融合の促進に向けたサイエンスマッピング
3. 学会等名 電気学会システム研究会 (招待講演)
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------