

令和 4 年 6 月 28 日現在

機関番号：82401

研究種目：若手研究

研究期間：2019～2021

課題番号：19K20231

研究課題名（和文）動的情報なしのネットワークに対する動的特徴を表す成長機構の推定方法の構築

研究課題名（英文）Estimating network growth mechanisms without network growth information

研究代表者

PHAM THONG (PHAM, THONG)

国立研究開発法人理化学研究所・革新知能統合研究センター・特別研究員

研究者番号：30803530

交付決定額（研究期間全体）：（直接経費） 2,800,000円

研究成果の概要（和文）：ユーチューブフォロワーや論文共著等のネットワークの特徴を分析するのによく使われているのは、確率ランダムネットワークモデルである。本研究は、確率ランダムネットワークモデルの推定方法とモデリング方法に貢献した。一つ目の貢献では、ネットワークの成長過程といった動的情報を観測できない時に優先的選択を推定できる方法を提案した。二つ目の貢献では、優先的選択と推移性を同時にモデリングするネットワーク成長モデル・そのモデルの推定方法を提案した。三つ目の貢献では、エッジの間の非独立性を許すハイパーグラフ成長モデルを提案し、そのモデルにおける推定を可能とした尤度の高速化計算方法も考案した。

研究成果の学術的意義や社会的意義

提案した諸手法・モデルによって、実世界におけるネットワークの性質に関する新たな知見を発見することが期待できる。まず、動的情報を観測できない時の優先的選択の推定方法によって成長履歴のないデータセットが初めて分析できるようになり、優先的選択に関する新たな発見に繋がることを期待できる。次に、優先的選択と推移性を融合した提案モデルを用いて、クラスタリング係数の高いネットワークの成長過程を分析でき、ネットワークの表現学習への貢献も期待できる。最後に、ハイパーグラフ成長モデルによって、今まで無視されてきたエッジの間の非独立性がモデル化され、エッジの間の複雑なパターンを発見できる可能性がある。

研究成果の概要（英文）：Growing complex networks have been used to model diverging real-world networks ranging from YouTube follower networks to academic collaborator networks. We have achieved three main contributions in modelling and estimating growing complex networks. In our first contribution, we proposed the first method that can estimate the non-parametric preferential attachment function from a network snapshot without any time information. This method corrects a previously unnoticed selection bias. In our second contribution, we proposed a novel network model that combines preferential attachment and transitivity in order to better explain the ubiquitous combination of a power-law degree distribution with a high clustering coefficient found in many real-world networks. In our final contribution, we built a novel hypergraph-based network growth model to explain the dependencies between edges in real-world networks. We also provided a fast method to compute the normalizing factor of the model.

研究分野：情報学

キーワード：優先的選択 ランダムネットワーク理論 複雑ネットワーク 選択バイアス

## 1. 研究開始当初の背景

YouTube フォロワや論文共著等、実ネットワークの動的な変化が、複数の成長機構の影響を受けると考えられる。例えば、ノード  $i$  と  $j$  の間に新しいリンクが張られる確率が、 $i$  と  $j$  のそれぞれの次数、 $i$  と  $j$  の間の共通隣接ノードの数等、様々な要素に依存すると考えられる。確率アクター・ベース・モデル (stochastic actor-based model) はこのように動的ネットワークの変化を複数の成長機構で捉え、このモデルにおける推定方法は RSiena という R パッケージに実装された。RSiena は、2017 年に INSNA William D Richards 賞を受賞した等、動的ネットワークの標準的な分析ツールとして高く評価され、様々な分野で広く使われている。しかし、RSiena は成長履歴を観測できるネットワークしか扱えないが、実際に成長履歴のないネットワークの方が多い。本研究は、RSiena が扱えない成長履歴のないネットワークに対する様々な成長機構を推定できるフレームワークの構築を目指す。

## 2. 研究の目的

成長履歴無しのネットワークに対して、確率アクター・ベース・モデルにおける様々な成長機構を推定できるフレームワークを構築し、そのフレームワークで様々な分野で今まで分析できなかった成長履歴無しのデータに対して分析し、新しい知見を発見することを目指す。

## 3. 研究の方法

まず、ネットワークが大きい時に、成長機構に関する統計量が定常な状態に達することを利用する。また、パラメトリックの関数形ではなく、ノンパラにすれば、情報を十分に持つ領域を推定できる。その成長機構の低領域に関する情報を十分に持つことに注目して、推定方法を構築できる。以下のような事項を進めてきた。

- **優先的選択の成長機構に関する推定方法を完成。** 次数の高いところでどうすればいいのか。例えば、次数  $k$  の出現時刻の結果を利用、隣り合うの次数をビンニング。
- **成長機構ありの実データで優先的選択の推定方法の有効性を検討。** 成長履歴ありの実ネットワークで、提案手法と成長履歴を利用する最尤推定方を比較した。その際、実ネットワークのサイズ、初期状態、成長履歴の特徴等に注目して、それらの要因が提案手法の推定結果の精度にどのように影響するのかを考察した。
- **成長履歴なしの実データにおける優先的選択に関する知見を探求。** 専門用語の共起ネットワークで言語学における優先的選択の現象や研究機関間の共同研究ネットワークで研究機関の間の優先的選択の現象等、この分野で関心が高いと思われるデータを分析することを検討した。
- **推移性の成長機構と優先的選択を同時に推定できるような方法を開発。** 前述のように、他の成長機構に対しても、 $c_k$  に相当する量をうまく近似できれば、提案手法の

アプローチを適用できる可能性がある。推移性の推定に提案手法を拡張することを検討した。

- 成長機構ありの実データで優先的選択と推移性の同時推定方法の有効性を検討及び成長履歴なしの実データで推移性に関する知見を探求
- どの成長機構であれば推定できるのか。成長機構の低領域と高領域に分けられる成長機構であれば、本アプローチを適用可能である。そのように分けられなくても、何らかの変換によって、その成長機構における順序が現れる、ということが出来るかを検証した。
- この時点に推定できる成長機構をソフトウェアに実装・公開する。提案手法を無料ソフトウェアとして **CRAN** に公開した。その際、使用方法と実データ分析の例を含む vignette と呼ばれる論文も執筆・公開し、一般人でも提案手法を簡単に使えるようにした。

#### 4. 研究成果

本研究は、確率ランダムネットワークモデルの推定方法とモデリング方法に貢献した。一つ目の貢献では、ネットワークの成長過程といった動的情報を観測できない時に優先的選択を推定できる方法を提案した。

二つ目の貢献では、優先的選択と推移性を同時にモデリングするネットワーク成長モデル・そのモデルの推定方法を提案した。

三つ目の貢献では、エッジの間の非独立性を許すハイパーグラフ成長モデルを提案し、そのモデルにおける推定を可能とした尤度の高速化計算方法も考案した。

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件/うち国際共著 2件/うちオープンアクセス 4件）

1. 著者名 Pham Thong, Sheridan Paul, Shimodaira Hidetoshi	4. 巻 9
2. 論文標題 Non-parametric estimation of the preferential attachment function from one network snapshot	5. 発行年 2021年
3. 雑誌名 Journal of Complex Networks	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/comnet/cnab024	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
1. 著者名 Inoue Masaaki, Pham Thong, Shimodaira Hidetoshi	4. 巻 10
2. 論文標題 A Hypergraph Approach for Estimating Growth Mechanisms of Complex Networks	5. 発行年 2022年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 35012 ~ 35025
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/access.2022.3143612	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 Inoue Masaaki *, Pham Thong * (co-first), Shimodaira Hidetoshi	4. 巻 14
2. 論文標題 Joint estimation of non-parametric transitivity and preferential attachment functions in scientific co-authorship networks	5. 発行年 2020年
3. 雑誌名 Journal of Informetrics	6. 最初と最後の頁 101042 ~ 101042
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.joi.2020.101042	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 Thong Pham, Paul Sheridan, Hidetoshi Shimodaira	4. 巻 92
2. 論文標題 PAFit: An R Package for the Non-Parametric Estimation of Preferential Attachment and Node Fitness in Temporal Complex Networks	5. 発行年 2020年
3. 雑誌名 Journal of Statistical Software	6. 最初と最後の頁 1-30
掲載論文のDOI（デジタルオブジェクト識別子） 10.18637/jss.v092.i03	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5件（うち招待講演 1件 / うち国際学会 1件）

1. 発表者名 Pham Thong
2. 発表標題 複雑ネットワークの成長過程を観測できない時の優先的選択関数の推定方法
3. 学会等名 2021年度統計関連学会連合大会
4. 発表年 2021年

1. 発表者名 井上 雅章, Pham Thong, 下平 英寿
2. 発表標題 任意のノード特徴量による成長機構をもつハイパーグラフモデル
3. 学会等名 2021年度統計関連学会連合大会
4. 発表年 2021年

1. 発表者名 Pham Thong
2. 発表標題 Estimating Preferential Attachment in Growing Networks
3. 学会等名 International Symposium on New Developments of Theories and Methodologies for Large Complex Data (招待講演)
4. 発表年 2021年

1. 発表者名 井上雅章, Pham Thong, 下平英寿
2. 発表標題 ハイパーグラフ成長モデルにおける優先的選択関数の推定
3. 学会等名 2019年度統計関連連合大会
4. 発表年 2019年

1. 発表者名 Masaaki Inoue, Thong Pham, Hidetoshi Shimodaira
2. 発表標題 Statistical Estimation of the Effects of First and Second Order Local Structures on Growth of Complex Networks
3. 学会等名 ACML 2019 Workshop on Statistics and Machine Learning Researchers in Japan (国際学会)
4. 発表年 2019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
---------	---------