

科学研究費助成事業 研究成果報告書

令和元年6月13日現在

機関番号：15101

研究種目：基盤研究(C) (一般)

研究期間：2014～2018

課題番号：26330275

研究課題名(和文) ネットワークの階層性を利用した遺伝子ネットワークの同定精度改善に関する研究

研究課題名(英文) A study of the Improvement of Reliabilities of Regulations using a Hierarchical Structure in a Genetic Network

研究代表者

木村 周平 (KIMURA, Shuhei)

鳥取大学・工学研究科・教授

研究者番号：20342777

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：推定された遺伝子ネットワークには一般に多くの間違いが含まれている。そのような間違いを減らすために本研究では「生物ネットワークは階層性を持つ」という性質を利用する。ネットワークの階層性の検知にはClausetらの提案したHRGモデルを利用する。遺伝子ネットワーク同定法の推定した相互作用が、推定した階層性と矛盾する場合、それらの信用度を低いと判断する。しかしそもそも同定したネットワークには間違いが多いため、そこから階層性を推定することは簡単ではない。そこで本研究ではブートストラップ法と組み合わせた同定法によって、多数のネットワークを推定し、それらの多くに矛盾しない階層性を抽出する方法を提案した。

研究成果の学術的意義や社会的意義

遺伝子ネットワーク同定は、がんなどの遺伝子を原因とする疾病に対する薬剤の標的遺伝子を見つけるためなどに利用可能と考えられている。本研究では遺伝子ネットワークの同定精度を改善するために、これまでに利用できなかった「ネットワークの階層性」という性質を利用する方法を開発した。提案手法は階層性だけでなく、これまで利用の難しかった複数の知識を利用することが可能と考えられ、さらなる同定精度の改善が期待できる。

研究成果の概要(英文)：Genetic network inference methods generally infer many erroneous regulations. To decrease the number of erroneous regulations, this study uses a priori knowledge that biochemical networks exhibit hierarchical structures. This study detects the hierarchical structure in the target network using a hierarchical random graph model proposed by Clauset and colleagues. When the regulations inferred by the inference method are inconsistent with the detected hierarchical structure, we can conclude that they are unreasonable. However, it is not always easy to detect the hierarchical structure in the target network because of the regulations erroneously inferred by the inference method. To obtain a reasonable hierarchical structure, this study first infers many genetic networks from the observed gene expression data by using a bootstrap-based method. We then extract a hierarchical structure from the inferred multiple genetic networks so that it is consistent with most of the networks.

研究分野：知能システム科学

キーワード：遺伝子ネットワーク 階層性 ブートストラップ法

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

生物は、その遺伝子が mRNA やタンパク質を通じて互いに制御しあうことで、生命機能を維持している。しかし現在のところ、どの遺伝子がどの遺伝子を制御しているかに関しては、あまり良く分かっていない。遺伝子ネットワーク同定とは、RNA-seq といった生化学実験技術によって得られた遺伝子発現量(遺伝子がどの程度働いているかを表す量)の時系列データから、遺伝子間の制御関係を推定することである。遺伝子ネットワークを同定することが出来れば、例えばガンをはじめとする遺伝子の病気に対する薬剤の標的遺伝子を見つけることなどが容易になると考えられている。その有用性から、遺伝子ネットワーク同定技術には、近年注目が集まっている。研究代表者を含めた多くの研究者が、遺伝子ネットワーク同定法を提案している。ところがこれらの手法によって同定された遺伝子ネットワークには、一般に多くの間違っただ相互作用が含まれるという問題があった。

2. 研究の目的

本研究では、(1)「新たな事前知識の利用」、(2)「モデルの簡略化」という2つのアイデアを用いて上記の問題の解決を目指す。以下ではそれぞれのアイデアと、その実現に際しての問題点について説明する。

(1) 新たな事前知識の利用

本研究では「遺伝子ネットワークには階層構造がある」という性質を新たに事前知識として利用することで遺伝子ネットワークの精度向上を目指す。ネットワークの階層構造とは、密に相互作用するノードがグループをなし、それらのグループが集まってさらに大きなグループを構成する、といった繰り返し構造のことである。本研究ではClausetらの提案したHierarchical Random Graphを用いる方法を利用して、同定された遺伝子ネットワークの階層構造を推定する。推定された階層構造を利用すれば、それに矛盾しない遺伝子間相互作用は信頼度が高く、矛盾する遺伝子間相互作用は信頼度が低いと判断することができる。

(2) モデルの簡略化

遺伝子ネットワークの同定に頻繁に利用されるモデルの一つとして、S-system と呼ばれる微分方程式モデルがある。S-system モデルに基づく遺伝子ネットワーク同定法は数多く提案されており、研究代表者らも過去に幾つかの同定法を提案している。しかし S-system モデルには、他の広く使われているモデルに比べて、パラメータ数が多いという問題がある。そのため S-system モデルに基づく遺伝子ネットワーク同定では多くの遺伝子発現データを与える必要があり、そのことが S-system モデルに基づく方法の欠点となっていた。これに対して近年 Chemmangattuvalappil らは、S-system モデルの一部のパラメータを定数に固定することで自由度を半減させたモデルを用いた遺伝子ネットワーク同定法を開発し、さらにそのようなモデルを用いても妥当な遺伝子ネットワークが同定可能であることを実験的に示した。しかし Chemmangattuvalappil らの同定方法には、計算コストが非常に高いという問題があった。

3. 研究の方法

(1) 簡略化した S-system モデルに基づく遺伝子ネットワークの高速同定法の開発

自由度を低減させた S-system モデルのパラメータ数は、 N をネットワークに含まれる遺伝子の数としたとき、 $N(N+2)$ 個となる。本研究ではこれらのパラメータ推定問題を、研究代表者の過去に開発した遺伝子ネットワーク同定法のアイデアを利用することで、 N 個の 1 次元関数最適化問題として定式化する。

(2) 遺伝子ネットワーク同定器へのブートストラップ法の適用

1年目に作成した遺伝子ネットワーク同定器とブートストラップ法を組み合わせる。

(3) ネットワークの階層構造推定法の開発

ここまでで開発した方法から出力される大量のネットワークから階層構造を推定する方法を開発する。そのために Clauset らの開発した方法を改良する予定である。

(4) 実験を通じた提案手法の有用性の確認

人工的に作成したネットワーク、及び実際のネットワークの解析を通して提案手法の有用性を確認する。提案手法はネットワークの階層性を考慮するため、その有用性を確認するためには比較的大規模なネットワークの解析が必要である。人工ネットワークには DREAM プロジェクトのデータを利用する予定である。DREAM プロジェクトとは生物系ネットワークの解析手法を競いながら開発する世界的な枠組みであり、そこで利用される人工データは現実のネットワーク構造を参考に作成されている。また実データは連携研究者の岡田(畠山)真里子氏(理化学研究所)から提供を受ける予定である。岡田氏とは10年以上にわたって共同研究を行ってきており、遺伝子ネットワーク解析の実績もある。

4. 研究成果

上記の方法(1)に関連して、高速に遺伝子ネットワークの簡略版 S-system モデルを同定する方法を開発した。その成果は雑誌論文(4)、国際会議論文(9)、及び学会発表(5)で発表を行った。

方法(2)に関しては、実際に開発した方法とブートストラップ法を組み合わせても、十分な同定精度を達成することは出来なかった。そのため本研究では大量の遺伝子ネットワークを出力する方法として、研究代表者が過去に提案した手法をそのまま利用した。

方法(3)と方法(4)に関しては実際に階層構造を利用して遺伝子ネットワークの同定精度を改善する方法を開発し、人工データ、及び実データを用いてその性能を確認した。その成果は雑誌論文(3)、国際会議論文(8)、及び学会発表(4)で発表した。

これまでの研究の過程で、必ずしも階層構造だけでなく、他の事前知識も利用することでさらなる同定精度改善の可能性が示唆された。そこで本研究ではさらに提案手法を進展させ、ネットワークの階層性だけでなく、クラスタ性やスケールフリー性なども考慮しながら同定精度の改善を目指す手法を開発した。その成果は雑誌論文(2)で発表した。

本研究で開発した方法は、研究代表者の過去に開発した遺伝子ネットワーク同定法のみを利用していった。しかし多数の遺伝子ネットワークを同定する方法や、遺伝子間相互作用に信頼度を割り当てる方法であれば、どのような方法でも適用可能なはずである。そのことを示すために遺伝子間相互作用に信頼度を割り当てる方法の開発も行った。具体的にはランダムフォレストを利用した因子ネットワーク同定法の開発などを行い、非常に良好な同定精度を達成した。その成果は雑誌論文(1)、国際会議論文(5),(6),(7)、図書(1)、及び学会発表(1),(2),(3)などで発表した。

5. 主な発表論文等

〔雑誌論文〕(計9件)

(1) S.Kimura, M.Tokuhisa and M.Okada: Inference of Genetic Networks using Random Forests: Assigning Different Weights for Gene Expression Data, Journal of Bioinformatics and Computational Biology, in press, DOI: 10.1142/S021972001950015X, 査読有。

(2) S.Kimura, K.Kitazawa, M.Tokuhisa and M.Okada-Hatakeyama: Using A Priori Knowledge after Genetic Network Inference: Integrating Multiple Kinds of Knowledge, Chem-Bio Informatics Journal, 17, 53-71, 2017, DOI: 10.1273/cbij.17.53, 査読有。

(3) S.Kimura, M.Tokuhisa and M.Okada-Hatakeyama: Genetic Network Inference using Hierarchical Structure, *Frontiers in Physiology*, 7, 57 (12 pages), 2016, DOI: 10.3389/fphys.2016.00057, 査読有.

(4) S.Kimura, M.Sato and M.Okada-Hatakeyama: An Effective Method for the Inference of Reduced S-system Models of Genetic Networks, *IP SJ Transactions on Bioinformatics*, 7, 30-38, 2014, DOI: 10.2197/ipsjtbio.7.30, 査読有.

(5) S.Kimura, M.Tokuhisa and M. Okada-Hatakeyama: Inference of Genetic Networks Using Random Forests: Use of Different Weights for Time-series and Static Gene Expression Data, *Proc. of the 18th IEEE International Conference on Bioinformatics and Bioengineering*, 98-103, 2018, DOI: 10.1109/BIBE.2018.00082, 査読有.

(6) S.Kimura, M.Tokuhisa and M. Okada-Hatakeyama: Inference of Genetic Networks from Time-series of Gene Expression Levels Using Random Forests, *Proc. of the 2017 Conference on Computational Intelligence in Bioinformatics and Computational Biology*, 22-27, 2017, 査読有.

(7) S.Kimura, M.Tokuhisa and M. Okada-Hatakeyama: Simultaneous Execution Method of Gene Clustering and Network Inference, *Proc. of the 2016 Conference on Computational Intelligence in Bioinformatics and Computational Biology*, 7 pages, 2016, 査読有.

(8) S.Kimura and M.Okada-Hatakeyama: Improvement of Reliabilities of Regulations using a Hierarchical Structure in a Genetic Network, *Proc. of the 2015 International Joint Conference on Neural Networks*, 485-491, 2015, 査読有.

(9) S.Kimura, M. Furuta, M. Sato and M. Okada-Hatakeyama: A Fast Parameter Estimation Method for Inferring Reduced S-system Models of Genetic Networks, *Proc. of 2014 International Conference on Information Systems and Computing Technology*, 45-50, 2014, 査読有.

〔学会発表〕(計 5 件)

(1) S.Kimura, M.Tokuhisa and M. Okada-Hatakeyama: Inference of Genetic Networks Using Random Forests: Use of Different Weights for Time-series and Static Gene Expression Data, *Proc. of the 18th IEEE International Conference on Bioinformatics and Bioengineering*, 98-103, 2018.

(2) S.Kimura, M.Tokuhisa and M. Okada-Hatakeyama: Inference of Genetic Networks from Time-series of Gene Expression Levels Using Random Forests, *Proc. of the 2017 Conference on Computational Intelligence in Bioinformatics and Computational Biology*, 22-27, 2017.

(3) S.Kimura, M.Tokuhisa and M. Okada-Hatakeyama: Simultaneous Execution Method of Gene Clustering and Network Inference, *Proc. of the 2016 Conference on Computational Intelligence in Bioinformatics and Computational Biology*, 7 pages, 2016.

(4) S.Kimura and M.Okada-Hatakeyama: Improvement of Reliabilities of Regulations using a Hierarchical Structure in a Genetic Network, *Proc. of the 2015 International Joint Conference on Neural Networks*, 485-491, 2015.

(5) S.Kimura, M. Furuta, M. Sato and M. Okada-Hatakeyama: A Fast Parameter Estimation Method for Inferring Reduced S-system Models of Genetic Networks, *Proc. of 2014 International Conference on Information Systems and Computing Technology*, 45-50, 2014.

〔図書〕(計 1 件)

(1) S.Kimura: Inference of Vohradsky's Models of Genetic Networks Using Evolutionary Algorithms, H.Iba and N.Noman, eds.: *Evolutionary Computation in Regulatory Network Research*, Wiley, 96-117, 2016, Book chapter.

6 . 研究組織

(2)研究協力者

研究協力者氏名：岡田 眞里子

ローマ字氏名：OKADA-HATAKEYAMA, mariko

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、研究者個人に帰属されます。