

科学研究費助成事業 研究成果報告書

平成 29 年 5 月 16 日現在

機関番号：11301

研究種目：基盤研究(B) (一般)

研究期間：2013～2016

課題番号：25280089

研究課題名(和文)非加算の体積とエントロピーによるコミュニティ抽出の確率的計算モデル設計理論

研究課題名(英文) Design Theory of Probabilistic Computational Models for Community Detections based on Non-Additive Volume and Entropy

研究代表者

田中 和之 (TANAKA, Kazuyuki)

東北大学・情報科学研究科・教授

研究者番号：80217017

交付決定額(研究期間全体)：(直接経費) 13,200,000円

研究成果の概要(和文)：本研究課題ではネットワーク構造に対して与えられたデータからの高性能の確率的コミュニティ検出計算モデルの提案に成功した。この計算モデルは確率伝搬法と期待値最大化アルゴリズムを用いて構成されており、人工的に生成されたネットワークおよび現実世界のネットワークに対して良好な結果が得られることを確認している。成果の一部は都市ネットワークにおける交通量予測モデル、インターネットを介したの同時刻の複数の動画からの同一人物検出モデルの提案へと展開している。更に提案した確率的コミュニティ検出計算モデルをモジュラリティという評価指標の視点から改良し、量子アニーリング法の提案へと発展させている。

研究成果の概要(英文)：We have proposed a new algorithm to detect the community structure in a network that utilizes both the network structure and vertex attribute data. The problem addressed in this paper is the detection of the community structure from the information of both the network structure and the vertex attribute data. The detection of the community structure in our method is achieved by using belief propagation and an EM algorithm. We numerically verified the performance of our method using computer-generated networks and real world networks. Actually, a part of our research results and aspects in complex networks are expanded to the traffic data reconstruction problem and on-demand color calibration system to track pedestrians across non overlapping fields of fixed camera view. Moreover, we have proposed a new quantum annealing method to achieve the optimization of modularity in community detection problems.

研究分野：確率的情報処理

キーワード：数理工学 機械学習 統計数学 情報統計力学 確率的情報処理

1. 研究開始当初の背景

近年、スマートフォンの普及、センサー工学の発達とともに大量のデータの Web を通じての発信・取得が可能となり、従来では想定できなかったようなビッグデータを個人レベルで取り扱うことのできる時代となってきた。そのような社会情勢のなかで、ビッグデータからの必要な情報の抽出技術の確立は急務として近年多くのコンピュータサイエンスの分野でデータマイニングとして研究が進められてきている。その中で特に重要な要素技術のひとつにコミュニティ抽出がある。コミュニティ抽出は大規模な基本構成要素が互いに関連しながら集まる構造のなかで、特に関連性の強い基本構成要素ごとにグルーピングを行うことを目的とする技術であり、その用途は人間関係のソーシャルネットワークはもとより、経済における企業、製品、原材料および消費者から構成されるサービスネットワーク、通信工学におけるアドホックネットワーク、交通工学における地理的ネットワークなど多岐にわたる。コミュニティ抽出の基本的アルゴリズムのいくつかは米国 Michigan 大学の M. E. J. Newman 教授等の研究グループにより提案されており、これらはネットワーク科学と計算理論を確率的情報処理へと展開するものとして注目されている。この確率モデルは確率的情報処理の分野では Potts モデルと呼ばれるものに近い構造を持ち、このことに着目した仏国 CEA Saclay の L. Zdeborova 研究員はより一般化された Potts モデルと確率伝搬法(Belief Propagation)にもとづくアルゴリズムを提案し、その有効性を Physical Review E, Vol.84, Article No.066106 (2011) 等において報告している。

2. 研究の目的

本計画研究ではコミュニティ抽出システムのビッグデータ解析への応用のための理論体系の基盤の構築を確率伝搬法のもとでのコミュニティのエントロピー最大化基準と非加算的体積最小化基準を組み合わせることで行うことを目的としている。コミュニティ抽出は頂点と辺からなるネットワークを複数のグループに分類する方法である。ビッグデータを生成するネットワークは空間構造が一樣ではなく、統計的学習・予測モデルをそのまま構成したのでは膨大な数の相互作用パラメータを伴うものとなる。ネットワークのコミュニティ抽出をそのネットワーク構造の普遍的本質をとらえながら行う理論を構築することで、階層化された統計的学習・予測システムの設計理論の創出へと展開することが期待された。

3. 研究の方法

本計画研究では確率伝搬法によるコミュニティ抽出の計算モデルをまず行っている。その上でより一般的なネットワークにおけるコミュニティ抽出という視点において計算

モデルの構築を進めた。本研究では確率伝搬法のもとでの計算モデルにおけるポテンシャル関数の関数形の教師有り学習による同定法に基づいて行っている。同定されたポテンシャル関数に対して研究代表者が確率的画像処理を目的として開発した拘束条件付きエントロピー最大化基準による確率伝搬法の計算モデルをコミュニティ抽出に転用する形でアルゴリズムの実装を進めた。その上で研究代表者が J. Phys. Soc. Jpn, 75, 115001 (2006)において提案した「非加算的体積最小化基準に基づくコミュニティ抽出法」と組合せることによる新しい計算モデルへの拡張し、実世界の具体的なネットワークシステムへの応用に向けて展開し、複合領域の拠点形成の核となり得る普遍的理論体系の基盤形成を目指した。

4. 研究成果

本計画研究の主な成果は以下の4点である。

- (1) 拘束条件付きエントロピー最大化により定式化されたマルコフ確率場の相転移現象を一次相転移の視点から解析し、状態数が増えることにより相転移現象の発現メカニズムの変化について明らかにした。その成果の一部はローマ大学ラ・サピエンツァ校物理学科の Federico Ricci-Tergenghi 准教授およびその研究グループとの国際共同研究へと発展しつつある。
- (2) 生成モデルとネットワークの統計的性質が本研究計画で提案した計算モデルに与える影響の分析を通して高性能の確率的コミュニティ検出計算モデルの構築に成功した。
- (3) 得られた成果をもとに成果の一部は都市ネットワークにおける交通量予測モデル、インターネットを介しての同時刻の複数の動画からの同一人物検出モデルの提案へと展開した。
- (4) コミュニティ検出モデルにおけるモジュラリティと呼ばれる評価関数に着目し、その最適化問題における量子アニーリング法の提案へと発展させている。

特に(4)は当初の研究計画では想定されなかった成果である。

以下、各年度ごとに成果の得られた経過について述べる。

平成 25 年度は一般化されたポテンシャル関数に基づく確率伝搬法を用いたコミュニティ抽出の計算モデルによる予測アルゴリズムの設計を進めた。従来法である最尤推定法で想定される問題は対数尤度がハイパラメータに関して微分不可能となる相転移が生ずる点にある。定式化した計算モデルから得られる対数尤度のハイパラメータに関する微分可能性について、標準データにおけるグラフ上での数値実験を通して詳細に解析を行ったが、ネットワーク構造を反映して、複数の相転移点が存在し、このことがコ

コミュニティ検出の予測アルゴリズムの動作の複雑な不安定性につながっていることを明らかにすることができた。このことは従来法による解析では確認することができなかったものであり、研究代表者の提案している高速条件付きエントロピー最大化に基づく期待値最大化アルゴリズムによるハイパラメータ推定の立場による解析を用いることで初めて明らかになった新しい知見であり、当初の研究計画立案の段階では想定されなかった研究成果である。平成 25 年度後期において University of Amsterdam の Max Welling 教授を海外研究協力者に加え、研究代表者が平成 26 年 3 月に同氏を訪問し、得られた研究成果の一部を説明し、次年度以降の研究計画に関する助言を得ている。また、研究分担者の安田宗樹准教授（山形大学）、海外研究協力者の Cyril Furtlehner 研究員（フランス国立情報学自動制御研究所）との協力により、研究成果の一部を都市ネットワークにおける交通量予測モデルへと展開する研究を行い、その成果をとりまとめた論文は学術雑誌 *Inverse Problems* に掲載され、最近の注目すべき論文の一つとして同学術雑誌の Webpage の Insights で紹介された。

平成 26 年度は確率伝搬法にもとづくコミュニティ抽出の計算モデルの数値実験、統計的解析を実施した。生成モデルとネットワークの統計的性質が提案した計算モデルに与える影響の分析を通して高性能の確率的コミュニティ検出計算モデルの構築に成功し、その成果の一部は国際会議 CIMCA2014 における査読付き論文として公開した。またネットワークの例えば次数分布、次数相関などに代表されるような統計的性質が尤度関数のハイパラメータすなわち相互作用パラメータの関数としての微分可能性に与える影響も解析も行っている。そこで得られた知見を画像の領域分割に転用し、確率伝搬法と拘束条件付きエントロピー最大化にもとづく計算モデルによる良好なモデル選択と画像領域分割が達成できることを確認した。その成果の一部は査読付き学術論文「*J. Phys. Soc. Jpn.*, Vol.83, Article ID.124002 (2014)」として公開した。更に研究代表者は研究分担者である安田宗樹准教授との協力により同計算モデルに統計物理学における伝統的計算技法のひとつである実空間繰り込み群法を組み合わせることにより計算速度の大幅な改善を行うことに成功した。これは当初の研究計画立案の段階では想定されなかった成果として評価され、その成果の一部は査読付き学術論文として「*J. Phys. Soc. Jpn.*, Vol.83, No.12, Article ID.124002」として公開している。更に研究分担者の安田宗樹准教授との協力により複雑ネットワーク上の計算モデルの統計的性能評価に関する定式化を進め、その成果の一部は査読付き原著論文「*IIS*, Vol.21, No.1, pp.11-23 (2015)」として公開した。また、確率伝搬法と線形応

答理論を組み合わせた新たなアルゴリズムの具体的設計を行った。

平成 27 年度は平成 26 年度に引き続いて確率伝搬法にもとづくコミュニティ抽出の計算モデルの数値実験、統計的解析を実施するとともに、その統計的機械学習モデルへと展開する研究を実施した。特に研究代表者は研究分担者である片岡駿助教と協力してコミュニティ検出においてもっとも重要な指標とされるモジュラリティそのものを用いた最適化に確率伝搬法にもとづく計算モデルを適用することで良好なモデル選択が達成できる可能性があることを数値実験を通して部分的に確認している。更に研究代表者は研究分担者である安田宗樹准教授との協力により、平成 26 年度に提案した実空間繰り込み群により計算モデルの高速化を波数空間繰り込み群法に展開し、計算速度の改善を行う研究を進め、その出発点としてガウシアングラフィカルモデルによる統計的機械学習モデルにおいて計算速度の大幅な改善が達成できることを確認した。これは当初の研究計画立案の段階では想定されなかった成果として評価される。更に研究分担者の安田宗樹准教授との協力により確率伝搬法を用いた複雑ネットワーク上の計算モデルの統計的性能評価法のレプリカ法を用いた高速化に関する定式化を進め、その成果の一部は査読付き原著論文「*Physical Review E*, Vol.92, No.4, Article ID.042120 (2015)」として公開した。また、研究代表者は研究分担者の片岡駿助教と協力して、本研究課題でこれまで得られた知見と計算モデルを都市における避難所割り当て問題に転用する研究へと発展させている。

平成 28 年度は平成 27 年度に引き続き確率伝搬法にもとづくコミュニティ抽出の計算モデルの数値実験、統計的解析を実施した。生成モデルとネットワークの統計的性質が提案した計算モデルに与える影響の分析を通して高性能の確率的コミュニティ検出計算モデルの構築に成功した平成 27 年度までの成果をさらに発展させ「*Journal of the Physical Society of Japan*, Vol.85, No.11, Article ID.114802 (2016)」における査読付き学術原著論文として公開した。更に研究代表者は山形大学の安田宗樹准教授との協力により同計算モデルに統計物理学における適応 TAP 自由エネルギーの摂動論的解釈を与えることに成功し、「*Journal of the Physical Society of Japan*, Vol.85, No.7, Article ID.075001 (2016)」における査読付き学術原著論文として公開した。また、日本大学の和泉勇治准教授と協力し、インターネットを介しての同時刻に撮影した複数の動画からの同一人物検出において本研究課題の研究成果の一部を転用する形での展開に成功し、その成果は「*IEEE Internet of Things Journal* Vol.4, No.2, pp.320-329 (2017)」として公開している。これらは当初の研究計画立案の

段階では想定されなかった成果として評価される。更にコミュニティ検出におけるモジュラリティを評価関数とする最適化問題に量子力学的揺らぎを導入した量子アニーリング法の提案を行うという、当初の研究計画では想定しなかった方向に展開している。その成果の一部は日本物理学会第 72 回年次大会において公開している。

研究計画はほぼ順調に進展し、その研究成果は主として 7 件の査読付き原著論文および 1 件の査読付き国際会議論文として公開している。一次相転移という確率モデルに内在する複雑かつ膨大な準安定状態の発現メカニズムに切り込みつつ、複雑ネットワーク上のコミュニティ検出の確率的計算モデルの構築という本計画研究の最大の課題を達成し、更に計算速度の改善という目的に向けて実空間繰り込み群法という統計物理学の伝統的手法をこれまでにない発想で統計的機械学習理論に世界で初めて導入することに成功した。更に、繰り込み群理論による統計的機械学習アルゴリズムの高速化と量子力学的揺らぎの導入によるクラスタリング手法の高速化に応用した研究へと発展させた成果は計画立案段階には全く想定されなかった成果である。以上の理由により当初の目的を十分に達成できたものと判断している。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 12 件)

Yuji Waizumi, Masako Omachi and Kazuyuki Tanaka: On Demand Color Calibration for Pedestrian Tracking in Non-overlapping Fields of View, IEEE Internet of Things Journal, Vol.4, No.2, April 2017, pp.320-329 (査読有, DOI:10.1109/JIOT.2016.2557814).

Shun Kataoka, Takuto Kobayashi, Muneki Yasuda and Kazuyuki Tanaka: Community Detection Algorithm Combining Stochastic Block Model and Attribute Data Clustering, Journal of the Physical Society of Japan, Vol.85, No.11, November 2016, Article ID.114802, pp.1-10 (査読有, DOI:10.7566/JPSJ.85.114802).

Muneki Yasuda, Chako Takahashi and Kazuyuki Tanaka: Perturbative Interpretation of Adaptive Thouless-Anderson-Palmer Free Energy, Journal of the Physical Society of Japan, Vol.85, No.7, July 2016, Article ID.075001, pp.1-2 (査読有, DOI:10.7566/JPSJ.85.075001).

古市智大, 片岡駿, 田中和之: 確率伝搬法による画像の領域分割における超事前分布を用いたパラメータ推定, 電

子情報通信学会信学技報, Vol.115, No.384, December 2015, pp.37-42 (査読無).

吉田健人, 片岡駿, 田中和之: 確率伝搬法と階調変換を用いた画像のノイズ除去アルゴリズムの高速化, 電子情報通信学会信学技報, Vol.115, No.384, December 2015, pp.31-36 (査読無).

真柴隆一, 片岡駿, 田中和之: 非欠損画素の情報及び確率伝搬法を用いた欠損画素の推定, 電子情報通信学会信学技報, Vol.115, No.384, December 2015, pp.13-188 (査読無).

Muneki Yasuda, Shun Kataoka and Kazuyuki Tanaka: Statistical analysis of loopy belief propagation in random fields, Physical Review E, Vol.92, No.4, October 2015, Article ID.042120, pp1-15 (査読有, DOI:10.1103/PhysRevE.92.042120).

Kazuyuki Tanaka, Shun Kataoka, Muneki Yasuda and Masayuki Ohzeki: Inverse Renormalization Group Transformation in Bayesian Image Segmentations, Journal of the Physical Society of Japan, Vol.84, No.4, April 2015, Article ID.045001, pp.1-2 (査読有, DOI:10.7566/JPSJ.84.045001).

斎藤拓馬, 和泉勇治, 田中和之: CNN と SVM を組み合わせた歩行者検知の効率化に関する一検討, 電子情報通信学会信学技報 Vol.114, Nos.520-521, March 2015, pp.41-46 (査読無).

Kazuyuki Tanaka, Shun Kataoka, Muneki Yasuda, Yuji Waizumi and Chiou-Ting Hsu: Bayesian Image Segmentations by Potts Prior and Loopy Belief Propagation, Journal of the Physical Society of Japan, Vol.83, No.12, December 2014, Article ID.124002, pp.1-11 (査読有, DOI:10.7566/JPSJ.83.124002).

Shun Kataoka, Muneki Yasuda, Cyril Furtlehner and Kazuyuki Tanaka: Traffic Data Reconstruction based on Markov Random Field Modeling, Inverse Problems, Vol.30, No.2, February 2014, Article ID.025003, pp.1-14 (査読有, DOI:10.1088/0266-5611/30/2/025003).

内田直樹, 和泉勇治, 田中和之: 周辺情報の統計的評価による風景中の文字認識, 電子情報通信学会信学技報 Vol.113 No.346, December 2013, pp.11-16 (査読無).

[学会発表](計 2 件)

関優也, 片岡駿, 田中和之: 量子アニーリングを用いたコミュニティ検出, 日本物理学会第 72 回年次大会, 2017 年 3 月 17 日, 大阪大学豊中キャンパス(査読無, 講演番号:17aB11-1)

Shun Kataoka, Takuto Kobayashi,
Muneki Yasuda and Kazuyuki Tanaka:
Community Detection based on Markov
Random Field using Node Attributes,
Proceedings of 2014 International
Conference on Computational
Intelligence for Modelling, Control &
Automation Intelligent Agents, Web
Technologies & Internet Commerce
Innovation in Software Engineering
(CIMCA-IAWTIC'14), CD-ROM, (Online
Conference), pp.38-43, 2014 (査読有).

〔図書〕(計1件)

田中和之(分担執筆): 確率的グラフィ
カルモデル(鈴木讓, 植野真臣編), 第8
章 マルコフ確率場と確率的画像処理,
pp.195-228, 共立出版, July 2016
(ISBN978-4-320-11139-4).

〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

〔その他〕

ホームページ等

6. 研究組織

(1) 研究代表者

田中 和之 (TANAKA, KAZUYUKI)
東北大学・大学院情報科学研究科・教授
研究者番号: 80217017

(2) 研究分担者

安田 宗樹 (YASUDA, MUNEKI)
山形大学・大学院理工学研究科・准教授
研究者番号: 20532774

和泉 勇治 (WAIZUMI, YUJI)

日本大学・工学部・准教授
研究者番号: 90333872

片岡 駿 (KATAOKA, SHUN)

東北大学・大学院情報科学研究科・助教
研究者番号: 50737278

(3) 連携研究者

木下 賢吾 (KINOSHITA, KENGO)
東北大学・大学院情報科学研究科・教授
研究者番号: 60332293

(4) 研究協力者

Chiou-Ting Hsu
台湾国立清華大学・コンピュータ科学部・教
授

Cyril Furtlehner

フランス国立情報学自動制御研究所・研究員

Lenka Zdeboraba

サクレー原子力センター・研究員

Federico Ricci-Tersenghi

ローマ大学ラ・サピエンツァ校・物理学科・
准教授

Max Welling

アムステルダム大学・理学部・教授

Pan Zhang

中国科学院理論物理研究所・准教授