

令和 5 年 6 月 16 日現在

機関番号：14602

研究種目：基盤研究(C) (一般)

研究期間：2018～2022

課題番号：18K11436

研究課題名(和文) 線グラフに基づくネットワークからの多重コミュニティ発見

研究課題名(英文) Overlapping Community discovery based on Line Graph

研究代表者

吉田 哲也 (Yoshida, Tetsuya)

奈良女子大学・工学系・教授

研究者番号：80294164

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：データの量や種類の増加に伴い、多様な価値観に応じてネットワーク上の資源を効果的に活用するための技術の確立が求められている。そこで、本研究課題では、線グラフに基づくネットワークからの多重コミュニティ発見に取り組んだ。具体的には、まずネットワークの接続関係を線グラフとして表現し、さらに領域知識としての関係の強さもネットワーク上の酔歩に基づいて線グラフ上の辺の重みとして統一的に表現することにより、線グラフに基づく多重コミュニティ発見法を定式化した。次に、定式化に基づく最適化学習アルゴリズムを開発し、開発したアルゴリズムを計算機システムとして実装するとともに、開発した手法を実データに適用して評価した。

研究成果の学術的意義や社会的意義

従来のアプローチではノードをそれぞれコミュニティに割り当てるノード分割を行うため、各ノードはひとつのコミュニティにしか所属できないという課題があった。しかし、実世界のネットワークにおいては、ノードに対応するユーザは関心や興味に応じて複数のコミュニティに所属することも多い。そこで、本研究ではノードやメッセージなどの内容ではなくネットワークの構造に基づいてリンクをコミュニティに割り当て、リンクの端点となるノードは接続するリンクが割り当てられたコミュニティにそれぞれ所属するとみなすことにより、ネットワークのリンク分割に基づく多重コミュニティ発見の実現に取り組んだ。

研究成果の概要(英文)：In order to cope with increasing quantity and variety of data, it is important to develop information technology which enables effective use of network data. We have developed a method for overlapping community discovery based on the line graph in graph theory. In the developed method, the relationship among nodes is represented as a line graph in graph theory, and the weights of the line graph is defined based on the weights in the original network. Under the framework of optimization learning, we have developed a overlapping community discovery algorithm based on the representation matrix of the given network. The algorithm has been implemented as a prototype system, and experiments over the prototype system were conducted over several real-world datasets. The results indicate the effectiveness of the developed learning method.

研究分野：知能情報学

キーワード：情報工学 機械学習 社会ネットワーク分析 コミュニティ発見

1. 研究開始当初の背景

(1) 高速なネットワークと廉価な大容量記憶装置の普及により、技術的な側面のみならず社会的な側面からもネットワークを活用する技術が注目を集めている。検索エンジンを通じた情報収集に加えて、blog や SNS などを通じた情報発信や、新たな人脈や機会を開拓するためのコミュニケーションの場としても広く活用され、欠かせない生活基盤となっている。ネットワークからの知識発見に関する研究はこれまで主に社会科学の分野で進められてきたが、スモール・ワールド現象を契機として、統計物理やランダムグラフ理論の観点からも様々な研究が始まった。個人などの対象を表すノード（頂点）とノードの関係を表すリンク（辺）から構成されるネットワークからのコミュニティ発見に対しては、コミュニティを対象（データ）の集まりが構成するクラスターと解釈することにより、ネットワーク（グラフ）に対するクラスタリングなどの機械学習の観点からも様々な研究が行われてきた。

(2) 従来のアプローチはコミュニティ発見を密な部分ネットワークを同定する問題と捉え、ノードをそれぞれコミュニティに割り当てるノード分割を行うため、各ノードはひとつのコミュニティにしか所属できないという課題があった。しかし、実世界のネットワークにおいては、ノードに対応するユーザは音楽や映画等の関心や興味に応じて複数のコミュニティに所属することも多いため、これを実現する技術の確立が求められている。ノードのプロファイルやリンクで伝達されるメッセージのラベルに基づき興味や関心などを扱うアプローチも考えられるが、ラベルに応じた分類処理などが更に必要になるという課題があった。

2. 研究の目的

(1) 本研究では、ネットワークの構造に基づいた多重コミュニティ発見を実現するために、線グラフに基づく多重コミュニティ発見法の研究開発を行う。単純グラフに対する線グラフとは、もとのグラフの辺を頂点とし、グラフで隣接する辺に対応する頂点を結んで得られるグラフであり、5 頂点以上ならば元のグラフの構造を線グラフから復元できることが知られている。

(2) 線グラフは構造の対応関係を保証しながらリンクとして表現される接続関係を明示的な対象として扱えるという利点がある。しかし、ネットワークでは一般に領域知識としての関係の強さなどが辺の重みとして表現されることが多いが、従来の線グラフではこの情報を活用できないという課題がある。そこで、本研究では接続関係（構造）に加えて関係の強さに着目し、リンク分割のためにネットワークの構造と辺の重みを統一的に重み付き線グラフとして表現して活用する多重コミュニティ発見の研究開発を行う。

3. 研究の方法

本研究では、ネットワークの構造に基づいた多重コミュニティ発見を実現するために、線グラフに基づく多重コミュニティ発見法の研究開発を目指した。具体的には、下記の3項目に対する研究開発を行った。

- ・線グラフに基づく多重コミュニティ発見の定式化
- ・最適化に基づく多重コミュニティ発見アルゴリズムの開発
- ・システム実装および実データでの評価・検証

4. 研究成果

(1) 研究代表者が従来から研究を進めてきたグラフ構造に基づく学習法を発展させて、まず、与えられたネットワークの構造をグラフ理論における線グラフに基づいて表現した。具体的には、接続行列に基づいて線グラフの構造を定義するとともに、ネットワーク上の酔歩に基づいて線グラフにおける辺の重みを定義した。さらに、ネットワーク上の酔歩で生成されてしまう自己ループを削除し、その重みを他の辺に分配することにより、ネットワークの構造と重みの保存を保証する重み付き線グラフを定義した。

(2) コミュニティ発見を(1)で定義する線グラフ上でのラベル学習とみなすことにより、多重コミュニティ発見をネットワークのリンク分割に基づいて定式化した。さらに、社会ネットワーク分析におけるモジュラリティ指標を多重コミュニティに拡張するとともに、ラベルを線グラフの辺に沿って伝播した際の定常解をネットワークからの学習と解釈することにより、多重コミュニティ発見を最適化問題として定式化した。

(3) 上記で定式化した目的関数の最適化を通じて多重コミュニティ発見を実現するために、行列演算に基づく最適化学習アルゴリズムを開発した。具体的には、コミュニティを表すラベルを連続変数に緩和することで、(2)で定式化した最適化問題に対する近似問題を定式化した。次に、グラフの表現行列の固有値分解に基づくスペクトルグラフ理論を活用することにより、重み

付き線グラフの表現行列に対する固有ベクトルを用いることで、リンク分割を通じて多重コミュニティ発見を実現するアルゴリズムを開発した。

(4) 上記で開発したアルゴリズムをプロトタイプシステムとして計算機上に実装した。また、実装したプロトタイプシステムの評価のために人工データを生成し、開発するアルゴリズムの実行速度やメモリ容量のスケーラビリティなどを評価した。さらに、ネットワークからの知識発見における代表的なベンチマークデータを収集して整備するとともに、収集したデータセットに開発したシステムを適用し、他手法との比較実験を行って開発手法の有効性を確認した。

(5) 本研究では、ネットワークの構造に基づいてリンクをコミュニティに割り当て、リンクの端点となるノードは接続するリンクが割り当てられたコミュニティにそれぞれ所属するとみなすことにより、ネットワークのリンク分割に基づく多重コミュニティ発見に取り組んだ。また、社会ネットワーク分析におけるモジュラリティ指標を多重コミュニティに拡張することで、多重コミュニティ発見の精度を従来の指標を素直に拡張して評価することに取り組んだ。実世界のネットワークにおいては、ノードに対応するユーザは関心や興味に応じて複数のコミュニティに所属することも多いため、本研究は、多様な価値観に応じて複数のコミュニティへの所属を可能とすることにつながると期待される。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 野村 涼子, 吉田 哲也	4. 巻 15
2. 論文標題 呼吸の特徴量を用いた心拍感覚の欠損補完	5. 発行年 2022年
3. 雑誌名 情報処理学会論文誌：数理モデル化と応用	6. 最初と最後の頁 11-18 (3)
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 オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Nomura, R. and Yoshida, T.	4. 巻 11
2. 論文標題 A Missing RR Intervals Complement Method Based on Respiratory Features	5. 発行年 2022年
3. 雑誌名 Journal of Advanced Biomedical Engineering	6. 最初と最後の頁 237-248
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 オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 0件）

〔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
---------	---------