

量子アニーリングを用いたコミュニティ検出 Community detection with quantum annealing

関 優也^{†‡} 片岡 駿[†] 田中 和之[†]
Yuya Seki Shun Kataoka Kazuyuki Tanaka

1. 概要

コミュニティ構造は複雑ネットワークを理解する上で重要な性質の1つである。コミュニティはネットワーク上の密に連結されたノードの集合として表される。密に連結されたノードを同じコミュニティに分類し、またより少ないエッジで繋がれたノードは別のコミュニティに分類する問題をコミュニティ検出と呼ぶ。分類の良さを表す指標の1つが Newman と Girvan によって提案されたモジュラリティ [1] である。モジュラリティは2つの項から成り、1つ目の項は同じコミュニティに属するノードをつなぐエッジの総数を表し、2つ目の項はコミュニティ構造が無いネットワークにおいて1項目と同じ量を評価したものである。1項目から2項目を差し引くことにより真にコミュニティ構造が存在する場合に大きな値を取る量を定義することができる。このようにして、コミュニティ検出はモジュラリティを最大化する最適化問題として定式化される。また、この問題はノードをコミュニティに分類する組み合わせに対してモジュラリティを最大化する問題なので、組み合わせ最適化問題の一種と考えられる。

近年、物理の分野で発展した手法を情報科学分野の問題に適用する研究が盛んに行なわれている。特に本研究と深い関わりがあるのが、組合せ最適化問題を解くための量子計算モデルである量子アニーリングである [2]。量子アニーリングは従来のアルゴリズムとは本質的に異なる原理で動作する計算モデルであり、量子性を用いることで従来のアルゴリズムでは指数関数時間かかる問題が多項式時間で解けることや、より少ない消費電力で問題を解くことが期待されている。しかしながら、量子アニーリングに関する研究は発展途上にあり、どのような問題に対して量子アニーリングが有効かという問題は未解決である。そこで本研究ではコミュニティ検出に対して量子アニーリングを適用し、その性能評価を行う。

量子アニーリングの有用性を判別する上で問題

になるのはその評価方法である。量子アニーリングを実行できる現存する量子計算機である D-Wave 計算機 [3] は、ハードウェア上の制約により今回の問題には使用できない。また、従来の計算機上で量子アニーリングのシミュレーションを原理通りに行う場合、ノード数に対して指数関数的なメモリサイズを必要とする。そこで、本研究では量子モンテカルロ法を用いて量子アニーリングの性能評価を行う。量子モンテカルロ法は量子アニーリングとは異なる原理で動作するアルゴリズムではあるが、量子アニーリングの性質をよく反映しており、また前述したメモリサイズの問題もない手法である。

2. モデル

この節ではモジュラリティの定式化を行う。本研究では次のコスト関数を用いる:

$$H(\{\sigma\}) = - \sum_{i \neq j} (A_{ij} - p_{ij}) \delta(\sigma_i, \sigma_j). \quad (1)$$

ここで、 A_{ij} は隣接行列、 p_{ij} はランダムグラフを仮定した際のノード i とノード j 間のエッジで結ばれる確率、 $\delta(\cdot, \cdot)$ は Kronecker のデルタ関数、 $\sigma_i \in \{1, \dots, q\}$ はノード i が所属するコミュニティのインデックスを表す。式 (1) の右辺の和は互いに異なる N 個全てのノードの組み合わせに関してとる。

コスト関数を定めるためには、式 (1) 中の p_{ij} を決める必要がある。今回はモジュラリティを再現する次の量を採用する:

$$p_{ij} = \frac{k_i k_j}{2M}. \quad (2)$$

ここで、 k_i ($i = 1, \dots, N$) はノード i を端点を持つエッジの数、 M は全体のエッジの総数を表す。また、本研究では量子モンテカルロ法を用いる際の便宜上の理由で、通常モジュラリティに $-2M$ を乗じた式をコスト関数として用いる。

[†] 東北大学大学院情報科学研究科

[‡] JST CREST

3. 手法

この節では量子モンテカルロ法を簡潔に導入する。量子モンテカルロ法では上述したモデルに仮想的な温度 T と横磁場と呼ばれる量子性を制御するパラメータ Γ を導入し、それらを制御しながらマルコフ連鎖モンテカルロ法を用いてコスト関数 (1) の最小値を探索する。従来のマルコフ連鎖モンテカルロ法は量子系に適用することができないが、鈴木-Trotter 分解 [4] を用いて量子系を対応する古典系に変換することで従来の手法を適用することが可能となる。この変換は元のグラフのコピーを M 個作成し、コピーされたグラフ間に新たなエッジを追加したグラフを生成する。変換後のコスト関数は次の式で与えられる:

$$\begin{aligned} H'(\{\sigma\}) &= - \sum_{\alpha=1}^M \sum_{i \neq j} \left(A_{ij} - \frac{k_i k_j}{2M} \right) \delta(\sigma_i(\alpha), \sigma_j(\alpha)) \\ &\quad + TM \ln \frac{\Gamma}{TM} \sum_{\alpha=1}^M \sum_i \delta(\sigma_i(\alpha), \sigma_i(\alpha+1)). \end{aligned} \quad (3)$$

量子モンテカルロ法を用いた手法の手順を次に示す。

1. T および Γ を大きい値に設定する
2. 温度 T を減少させつつ、コスト関数 (3) のもとでのマルコフ連鎖モンテカルロ法により状態を更新する
3. 温度を十分下げた後、横磁場 Γ を減少させつつ上と同様に状態を更新する
4. 横磁場を十分減少させた後、 M 個のコピーのうち最もモジュラリティの大きいコミュニティのインデックスのリストを解として出力する

4. 結果

この節では得られた結果の 1 つを示す。32 個のノードを持つコミュニティが 4 つあるグラフを考える。1 つのノードの持つエッジの数の期待値が 16 個になるようにし、そのうち 13 個が同じコミュニティに属するノードとつながり、残りの 3 つが異なるコミュニティに属するノードにつながるように調整する。図 1 は計算を始める前のコミュニティのインデックスを表しており、全てのノードが一つのコミュニティに属している。図 2 は量子モンテカルロ法適用後のコミュニティのインデックスを表す。密につながったノードが同じコミュニ

ティに属していることが見て取れ、量子モンテカルロ法により解が得られていることがわかる。

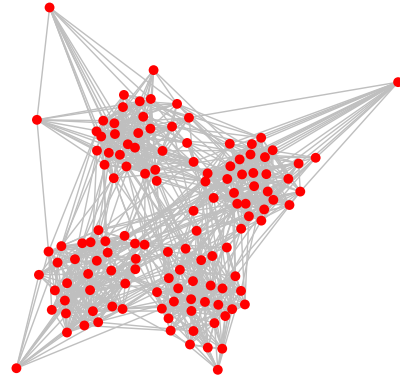


図 1 手法を適用する前のコミュニティのインデックス

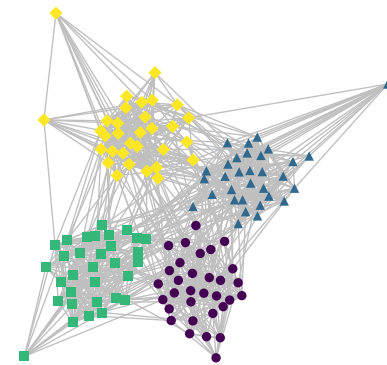


図 2 手法適用後のコミュニティのインデックス。同じ形のノードは同じコミュニティに属する。

5. 結論

量子モンテカルロ法を用いたモジュラリティ最大化によるコミュニティ検出を行い、実際にコミュニティ検出が可能であることを確かめた。この結果は、コミュニティ検出に量子計算が有効である可能性を示唆している。

謝辞

本研究は JST CREST 「ビックデータ時代に向けた革新的アルゴリズム基盤」の支援を受けたものである。

参考文献

- [1] M. E. J. Newman, and M. Girvan, Phys. Rev. E, **69**, 026113 (2004).
- [2] T. Kadowaki, and H. Nishimori, Phys. Rev. E, **58**, 5355 (1998).
- [3] M. W. Johnson *et al.*, Nature **473**, 194 (2011).
- [4] M. Suzuki, Prog. Theor. Phys., **56**, 1454 (1976).