

コミュニティ構造の影響を考慮したネットワーク成長モデルの提案

三好栄次†

鈴木育男†

山本雅人†

古川正志†

†北海道大学 大学院情報科学研究科

1 はじめに

複雑ネットワークに関する研究の一分野として、ネットワークを構成している各要素間の結合構造に着目し、ネットワークの特徴を把握しようとする研究が数多くなされている。その多くは、ネットワーク中の特徴量によりネットワーク成長を行うことに主眼が置かれている。それ以外にも複雑ネットワークの研究分野では、構成要素の局所的なルールに基づいたネットワークの成長モデルの研究も盛んに行われている。このネットワークの成長モデルの研究は、既存のネットワークの予測をする上で重要な研究である。しかし、現実のネットワークにおいて大規模なネットワークにおいてネットワーク全体の情報を把握できることはほぼ不可能である。この意味ではこれまでの成長モデルは、現実のネットワークの成長を模倣できているとは言い難い。

そこで、本研究では、既存のネットワーク成長モデルである CNN モデル [1] に対し、コミュニティ抽出を適用し、ネットワークの成長のルールにコミュニティ情報を付加した成長モデルの提案を行う。また、提案手法を適用したネットワーク生成による特徴量について、時間経過による変化を観測することにより、ネットワークの成長過程について構成と成長を考察する。

2 関連研究

2.1 Local Modularity

本研究で用いる、Clauset が提案した LocalModularity: R [2] を用いた局所コミュニティ抽出法について説明する。 V, E をノード集合及びリンク集合とするグラフ全体は $G = (V, E)$ となる。すでにコミュニティとなっている部分を $C (C \subseteq G)$ 、 C に含まれていないノード集合において、 C に含まれるノードに対しリンクがある部分集合を $U (U \subseteq G)$ 、 C に含まれているノードにおいて U に含まれるノードとリンクがある部分集合を $B (B \subseteq G)$ とする。 B に関して、隣接行列は以下の式で求められる。

$$b_{ij} = \begin{cases} 1: \text{ノード } v_i, v_j \text{ が繋がっており,} \\ v_i, v_j \text{ のどちらかは } B \text{ に属する} \\ 0: \text{それ以外} \end{cases} \quad (1)$$

Proposal of the network growth model by influence of community structure

†Eiji Miyoshi †Ikuo SUZUKI †Masahito YAMAMOTO †Masashi FURUKAWA

†Graduate School of Information Science and Technology, Hokkaido University

B の隣接行列から次の計算で LocalModularity: R が求まる。

$$R = \frac{\sum_{ij} b_{ij} \delta(i, j)}{\sum_{ij} b_{ij}} = \frac{I}{T} \quad (2)$$

T は B の総リンク数、 I は B のリンクにおいて U と繋がっていないリンクの数となる。

2.2 Connecting Nearest Neighbor(CNN) モデル

このモデルは「スケールフリー性」、「クラスタ性」、「スモールワールド性」を満たすソーシャルネットワーク的なグラフを生成するアルゴリズムの一つであり、「友達の友達に友達」の関係をポテンシャルリンクを用いて表現したネットワーク成長モデルである。 [1]

3 提案モデル

本研究では CNN モデルのネットワークの成長方法にコミュニティの概念を取り入れる。 CNN モデルにおいて、ポテンシャルリンクが発生する範囲を隣接ノードからコミュニティに広げることにより同好の集団とは知り合いになりやすくなる条件ことを表現する。

提案方法の局所コミュニティ抽出は、コミュニティ抽出の出発点とするノードから遠く離れたノードの情報は必要なくなる。 よって、ネットワーク全体のトポロジの影響を受けない。 また、同じノードで抽出されるコミュニティの範囲は、ネットワークが成長した場合でもノード付近のネットワークの構造が変わらない限り同じとなる。

ネットワークの生成方法を以下のように提案する。

1. 初期状態を決定する。
2. 以下の操作をノード数が一定数以上になるまで繰り返す。
 - (a) 確率 $1 - \mu$ でネットワーク中からランダムにノード v_i を選択する。
 - i. ノード v_i を核とした局所コミュニティを抽出する。
 - ii. ネットワークにノード v_j を追加し、リンクを張る。
 - iii. 抽出されたコミュニティに属するノードとノード v_j にポテンシャルリンクができる。

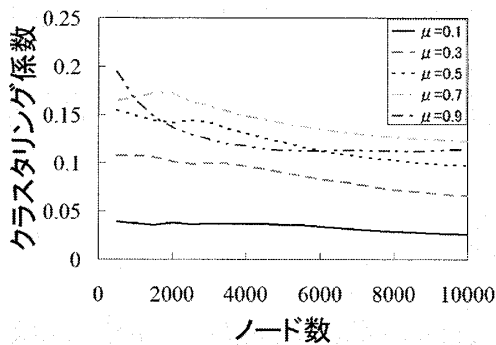


図 1: クラスタリング係数の推移

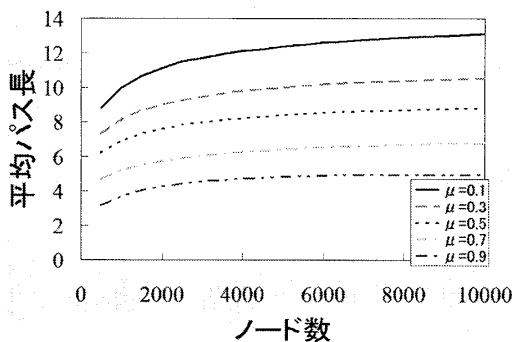


図 2: パス長の推移

(b) 確率 μ でポテンシャルリンクをランダムに 1 つ選び実際のリンクに変える。

今回の提案モデルでは、CNN モデルと比べて、ポテンシャルリンクの発生は 2 ノード先のノード間同士だけではないためクラスタリング係数の減少、ネットワーク中のリンクの増加に伴う局所コミュニティの抽出範囲拡大、局所コミュニティ抽出範囲拡大によるネットワーク直径、平均最短パスの減少、などが考えられる。

3.1 実験結果

実験結果を図 1, 図 2, 3 に示す。図 3 に関しては最も CNN モデルと異なった数値のみの図に載せた。

3.2 考察

図 1 より、クラスタリング係数の推移は低い値となっている。これは、CNN モデルと違い、2 次近傍のノードと、リンクを張っているためにクラスタリング係数が高くなっているのに対し、提案手法はポテンシャルリンクの存在範囲が 3 次近傍以上のノードとリンクが張られる可能性があるため、クラスタリング係数が低い。

図 2 は、ネットワークの平均パス長の推移である。ネットワークが初期の段階、ポテンシャルリンクの顕在化確率が高くなるにつれ、提案手法の平均パスは短くなる。これは提案手法においてより遠くのノードとリンクを張りやすくなるため、平均パス長が短くなる。

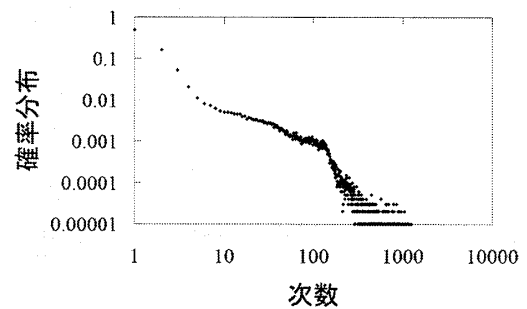
図 3: $\mu = 0.9$ における次数分布

図 3 は、顕在化確率を変更した場合のネットワークの最終状態における次数の確率分布である。顕在化確率が高くなるにつれて提案手法の確率分布はべき分布にはならず、次数がある値を超えると存在する確率が急に下がる。ネットワークにノードが追加される際、抽出されたコミュニティの多くはノード数 10 程度の小規模なコミュニティ構造となるためであると考えられる。今回のコミュニティ構造を考慮したモデルにおいては局所コミュニティによって顕在化確率が高くなると新規に追加されたノードは多くのポテンシャルリンクを持つようになる。

局所コミュニティの抽出方法の特徴により、ポテンシャルリンクの顕在化確率が高い場合において、図 3 で示すような次数がべき乗則にならない。

4 結言

本研究では、ネットワークの生成モデルに関して、以下のことを行った。

- コミュニティ構造を考慮したネットワーク構築のモデル化を提案した
- コミュニティ構造をネットワークの生成に用いることによりネットワーク中のリンク密度における、パラメータの変化を観測した。

本研究では実ネットワークを模したネットワークの成長モデルではなく、「コミュニティ構造」を定義、既存モデルに導入したが、現実のネットワークが持つ特徴を満たせていない部分が多く、課題としてパラメータ、成長ルールの追加およびコミュニティ抽出方法の変更などが挙げられる。

参考文献

- [1] Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations, May 2003.
- [2] Finding local community structure in networks. Phys. Rev. E, p. 026132, Aug 2005.