

Modularity を用いた複雑ネットワークの成長と構築に関する研究

三好栄次 † 鈴木育男 ‡

†北海道大学 工学部・情報工学コース

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

複雑ネットワークの成長モデルの研究が盛んに行われている。クラスタ構造や意図的に一部のノードにリンクが集中するようなネットワーク [1] や、ノード間でのつながりを強くするようなネットワーク [2] を生成する手法など提案されている。ネットワークの成長に関していろいろな生成ルールも存在する。

そこで本研究では、ネットワーク解析におけるコミュニティ分割手法を適用し、コミュニティの分割の良さを評価する Modularity を使い、組織内ネットワークでのノードの割り当て方法を提案し、各コミュニティ（部署）における人員の補強および同一部署でのグループ割り当てにおいてどのようなコミュニティ構造が形成されるかを示し、組織行動においてノードの特性によらない構成を目的としている。

2 関連研究

2.1 Modularity

Aaron Clauset [3] らが提案した方法であり、ネットワークの分割の良さを表す指標の Modularity である Q が高くなるような分割を目指すものである。Modularity はノード集合 V_l 内のノードの次数の総和とネットワークの全てのリンクとの割合を a_l ノード集合 V_l から V_m へのリンクの数とネットワーク中のリンクの総数 M の比を e_{lm} とすると

$$Q = \sum_{i \in 1 \dots l} Q_i = \sum_{i \in 1 \dots l} (e_{ii} - a_i^2) \quad (1)$$

となる。本研究では指標を Modularity とし Clauset らのアルゴリズムで抽出されたノード集合をコミュニティと呼ぶ。

2.2 コミュニティ分割アルゴリズム

Clauset らのアルゴリズムは Modularity の変化量 ΔQ を用い計算を高速化している。概要は以下のとおり

Research on growth and construction of complex network using with Modularity
†Eiji Miyoshi ‡Ikuo SUZUKI ‡Masahito YAMAMOTO ‡Masashi FURUKAWA
†Information Engineering, Hokkaido University
‡Graduate School of Information Science and Technology, Hokkaido University

1. 初期状態は 1 ノードずつ別のコミュニティとしコミュニティごとに結合した場合の Q の差分

$$Q_{(i+j)} = Q_i + Q_j + 2e_{ij} - 2a_i a_j \quad (2)$$

から $\Delta Q_{ij} = 2(e_{ij} - a_i a_j)$ を計算し最大のものを選ぶ。

2. ノード i, j が結合した場合ノード i, j と繋がっていたノード k に関して以下に従い Δ の更新をおこなう。

$$\Delta Q'_{jk} = \begin{cases} \Delta Q_{ik} + \Delta Q_{jk} : \text{if } k \text{ connected with } i, j \\ \Delta Q_{ik} - 2a_j a_k : \text{if } k \text{ connected with } i \\ \Delta Q_{jk} - 2a_i a_k : \text{if } k \text{ connected with } j \end{cases} \quad (3)$$

3. ΔQ がすべて負になるまで繰り返す

このとき得られる部分グラフはコミュニティ内でのリンク密度は高くコミュニティ外でのリンク密度は低くなるように結合されている。経験的にこの Modularity が 0.3~0.7 であればある程度まとまりのあるを形成しているといえる。

3 Modularity に基づくネットワークの成長モデル

本研究では Clauset[3] のコミュニティ分割を用い、成長モデルにおいて各ステップでのノードの追加し、同時にコミュニティの再定義を行い、まとまりの良いコミュニティ構造を形成させることを目的とした方法を以下に示す。

1. 初めに基になるネットワーク G を用意する
2. G をコミュニティ分割し、各コミュニティ i の Modularity: Q_i を計算する。
3. G にノード v_{new} を追加し、各コミュニティ i の Modularity: Q_i に比例した確率でコミュニティを選択する。同じコミュニティが複数回選ばれても良い。
4. 選択されたコミュニティ内から n 個ノード $\mathbf{V}_r = \{v_1, v_2, \dots, v_n\}$ を選択し、 v_{new} から \mathbf{V}_r へリンクを張る。
5. ノード数が N_{max} を超えるまで 2 から 4 を繰り返す

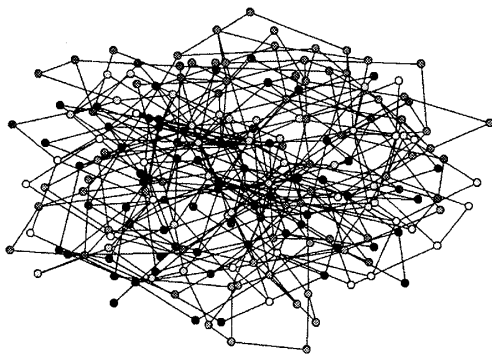


図 1: 200 ノードにより構成されたネットワーク

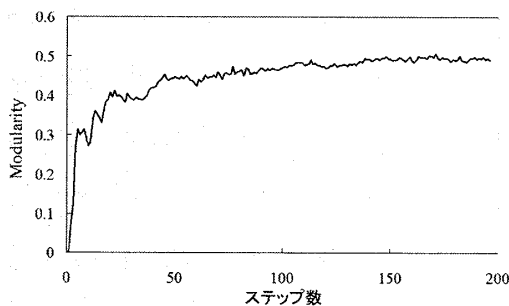


図 2: Modularity の推移

4 数値計算実験

4.1 実験条件

実験条件を以下に示す

1. G を 3 ノードによる完全グラフとする
2. パラメータとして $n=2$ $N_{max} = 200$ とする.
3. 得られたネットワークに対し, コミュニティ分割し, Modularity の推移, コミュニティサイズを観察する.

4.2 実験結果

図 1 に生成されたネットワーク, 図 2 に Modularity の推移をしめす. 図 2 より Modularity がある程度上昇したのち安定した状態となっていることがわかる. 全体を通して Modularity が 0.4 より大きくなっていることから生成されたネットワークにはコミュニティ構造が形成されていることがわかる

5 考察

ノードの追加確率は Modularity: Q_i によって決まる. Q_i は新しいノードを追加されると減少する機会が多いので各コミュニティごとの Modularity: Q_i の差があまりない可能性がある. このことから提案モデルでコミュニティ構造を持つネットワークが生成されることがわかる. 各コミュニティでのノード数の標準偏差をとったとこ

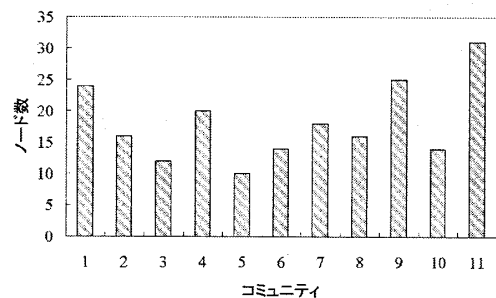


図 3: 最終状態におけるコミュニティごとのノード数

ろ 6.01 となり各コミュニティのノード数は最大 31, 最小 10 とノード数のばらつきが見られた. Modularity: Q_i によりコミュニティ i に張られるリンクの割合が決められるためである. したがってコミュニティ間でのノード数を均等にします.

6 結言

本研究ではネットワーク解析から得られるデータを基にネットワーク生成ルールを提案した. 具体的にはコミュニティを定義することで各コミュニティの Modularity: Q_i を得, ノードごとの特性を考慮せず Q_i にしたがってネットワークの生成ルールを提案した.

しかし, Modularity: Q_i のみでのノードの選択であるとコミュニティのノード数にばらつきが生じた. 一部のコミュニティにノードが集中する現象を企業における人事に例えたとの組織化においては効率が悪い場合がある. そのためコミュニティのノード数になるべく等しくなるような生成ルールも必要となる. さらに組織によっては部署数を一定とする場合もあり, 今後は, コミュニティ数固定および, 各コミュニティのノード数固定とした場合の Modularity の推移およびコミュニティ構造の形成の過程を調べる必要がある. また提案したモデルと実際人事管理システム等による組織構造と比較し, 提案モデルの妥当性の検証も行いたい.

参考文献

- [1] A.-L. Barabasi and R. Albert. Emergence of scaling in random networks, October 1999.
- [2] Alexei Vázquez. Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations, May 2003.
- [3] Aaron Clauset, M. E. J. Newman, and Cristopher Moore. Finding community structure in very large networks, Dec 2004.