

複雑ネットワークにおける ノードに内在する関係発生能力のモデル化と観測

矢野千紘[†] 武藤敦子[†] 森山甲一[†] 松井藤五郎^{††} 犬塚信博[†]

名古屋工業大学[†] 中部大学^{††}

1 はじめに

1960年代以降、社会学の分野において社会ネットワーク分析の研究が行われてきた。社会ネットワーク分析とは人や組織の持つ関係性をネットワークとして捉え、中心となる行為者や構築される派閥がどのようなものか解明する分野である。従来の中心性は、1時点のネットワークにおいて構造的に重要な頂点を評価するものであった。しかし、菅田らはネットワークの時系列において、辺の生成において重要な頂点を評価するものとして関係構築中心性 [1] を提案した。また、時田らは関係構築中心性を発展させ、2つの中心性が頂点に内在する能力とした。さらに関係構築中心性を用いて生成と観測のモデルとすること [2] を提案した。しかし時田らの手法では関係構築中心性の観測が難しかった。

本稿では同様の仮定において時田らの手法を改善し、関係構築中心性の観測を可能とするモデル化を目的とする。そこで関係構築中心性の再定義と新たなネットワーク変化手法を提案し、このモデルの下で数値を観測できることを示す。また、シミュレーション実験の結果を示す。

2 関連研究

本節では菅田らの関係構築中心性と、これを発展させた時田らのネットワーク変化のモデルを述べる。

2.1 関係構築中心性

菅田らは新しく関係が発生した2頂点 X, Y に関して X, Y の両方に関係を持っていた頂点 Z がある場合、 Z を関係発生 の 要因としてキューピッドと呼び、存在しない場合、 X, Y を関係発生 の 要因としてパイオニアと呼んだ。この考え方に基づいた中心性としてキューピッド値 (C 値)、パイオニア値 (P 値) を定義し、その値と頂点に関する他の属性との関係を調査した。

Modeling and Observation of the Inherent Ability of Nodes to Generate Relationships in Complex Networks

Chihiro Yano[†], Atsuko Mutoh[†], Koichi Moriyama[†], Tohgoroh Matsui^{††} and Nobuhiro Inuzuka[†]
Nagoya Institute of Technology[†], Chubu University^{††}

2.2 ネットワーク変化モデル

時田らは P 値、C 値を頂点に内在する特性と捉え直し、P 値を単位時間あたりに自身で関係を構築する数、C 値を隣接頂点の P 値に作用し影響を与える値と再定義した。二値が頂点に内在するとして、ネットワーク変化モデルを提案し、また2値をネットワークの変化から観測 (再現) する方法を検討した。

時田らのモデルでは各頂点 i が P 値 P_i 、C 値 C_i を持つことを仮定し、ある時刻で全ての頂点 i について以下の通り1ステップ変化する。

1. i に対して i と接続していない頂点 $j \in N - (NB(i) \cup \{i\})$ (以後、 N^{-i} と表記) をランダムに選ぶ。
2. i と j に共通の隣接頂点 k がひとつ以上ある場合、ランダムにひとつ k を選び確率 $P_i \times P_j \times C_k$ で辺を生成する。ない場合は確率 $P_i \times P_j$ で辺を生成する。
3. 手順2で辺が生成された場合、辺をランダムに減らす。ここで N はネットワークすべての頂点の集合、 $NB(i)$ は頂点 i の隣接頂点の集合である。

時田らのモデルでは、生成されたネットワークの変化から P 値、C 値を観測することができなかった。

3 ネットワーク変化モデルと観測法の提案

本研究ではネットワーク変化の生成とそこからの P 値、C 値の観測を可能とすることを目標として、菅田ら及び時田らの提案した関係構築中心性を再定義する。

3.1 関係構築中心性の再定義

本研究ではネットワークの変化について次のように再定義した。時刻は離散的で、1時刻において各頂点に辺を生成する機会を与える。頂点 i は P 値 P_i の確率で辺を生成する。各頂点が辺を生成する際、相手頂点の P 値に比例する確率でランダムに接続する。ただし、その頂点と共通する隣接頂点がある場合、その C 値を乗じた値を生成比率 $Link_i$ とする。共通する頂点が複数ある場合、すべての C 値を乗算する。

すなわち、頂点 i において、相手となる頂点 $j \in N^{-i}$ を次の $Link_i(j)$ に比例する確率で1つ選択する。

$$Link_i(j) = P_j \left(\prod_{k \in NB(i,j)} C_k \right)$$

ここで、 $NB(i, j)$ は $NB(i) \cap NB(j)$ を意味する。

最後にその時刻に生成された辺の数と同じ数の辺をネットワーク全体からランダムに削除する。

以上の手順をまとめる。提案するモデルはネットワーク $G = (N, E)$ の全ての頂点 $i \in N$ について次を行う。

1. 全ての $j \in N^{-i}$ について $Link_i(j)$ の比率で頂点 j を選択し、確率 P_i で辺 (i, j) を E に加える。
2. 手順1で加えた辺の数だけネットワーク全体からランダムに辺を削除する。

3.2 観測

P 値, C 値をもつネットワークが前節のモデルで変化したときに P 値を観測することを目的とする。

まず、単位時間で頂点 i に追加される辺数の期待値 = i が自ら生成する辺数 + 他のノードから生成された辺数の期待値である。第1項は、 P_i そのものである。そこで、第2項を見積もる必要がある。これは i へ接続が可能な頂点集合 N^{-i} が i を対象に選択する確率の和であり、以下の通りである。

$$e_i = \sum_{j \in N^{-i}} P_j \frac{Link_j(i)}{\sum_{x \in N^{-j}} Link_j(x)}$$

$$= \sum_{j \in N^{-i}} P_j \frac{P_i (\prod_{k \in NB(j,i)} C_k)}{\sum_{x \in N^{-j}} P_x (\prod_{k \in NB(j,x)} C_k)} \quad (\#)$$

align 環境で
= をそろえる

を e_i とする。 e_i

ここでは j の隣接頂点が多いとき、 i, j 間の共通頂点と x, j 間の共通頂点は多くなると予想される。少ない場合も同様であり、 i, j 間に影響する C 値と x, j 間に影響する C 値は同程度だと考え無視して次の通り近似する。

$$e_i \approx \sum_{j \in N^{-i}} P_j \frac{P_i}{\sum_{x \in N^{-j}} P_x} = P_i \sum_{j \in N^{-i}} P_j \frac{1}{\sum_{x \in N^{-j}} P_x}$$

各頂点 a について、 a と隣接しない頂点の P 値の合計 $\sum_{x \in N^{-a}} P_x$ は変化が小さいため定数 A で近似する。

$$e_i \approx P_i \sum_{j \in N^{-i}} P_j \frac{1}{A} = P_i \frac{1}{A} \sum_{j \in N^{-i}} P_j = P_i$$

したがって、頂点 i に対して他の頂点の関係発生能力によって生成される辺数の期待値も P_i と見積もられる。即ち、自ら生成する辺と他の頂点から接続される辺を合わせた生成される辺数の期待値は $2P_i$ である。

そこで、頂点 i について各時刻で増加する辺数の平均を観測し、その $1/2$ を i の P 値 P_i として観測する。

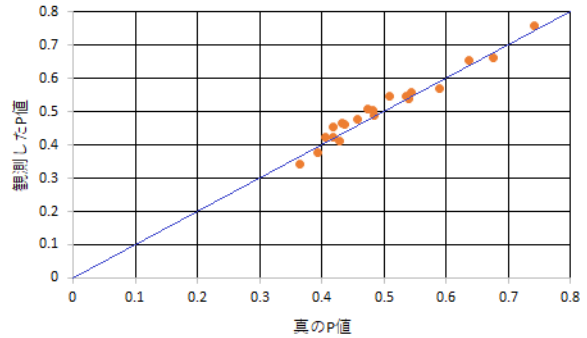
4 シミュレーション実験

頂点数 20 のランダムネットワークを提案モデルで 1000 回変化させることでネットワークの時系列を生成

し、その結果から観測モデルによって P 値の観測 (復元) を行った。

ネットワークの各頂点の P 値を平均 0.5, 分散 0.1 の正規分布で与え、C 値を 1 または 10 のランダム値とし、各頂点の P 値 (真の P 値) と観測した P 値の値を図 1 に示す。

図 1: P 値とその観測値



観測値は P 値に近い値になっており、提案手法での P 値の観測が可能なのが確認された。

また、表 1 に P 値, C 値の条件を変えて実験を行った場合の P 値と観測値の平均二乗誤差を示す。表より、本手法による P 値の観測は、P 値, C 値の値に大きく依存しないことが分かる。

P 値	C 値	二乗誤差
正規分布 $\mu: 0.5, \sigma: 0.1$	1or10 ランダム	0.0095
正規分布 $\mu: 0.5, \sigma: 0.2$	1or10 ランダム	0.0094
正規分布 $\mu: 0.5, \sigma: 0.05$	1or10 ランダム	0.0077
正規分布 $\mu: 0.5, \sigma: 0.1$	1or30 ランダム	0.0114
正規分布 $\mu: 0.5, \sigma: 0.1$	1or50 ランダム	0.0054
正規分布 $\mu: 0.5, \sigma: 0.1$	1~10 一様分布	0.0047

表 1: P 値のその観測値の平均二乗誤差

5 まとめ

ネットワークから関係構築中心性を観測するため、ネットワーク変化とそこからの観測のモデルを提案した。実験の結果、P 値を観測可能なことを示した。今後、C 値の観測法を与え、ネットワークの変化からその関係構築の能力を測るモデルを完成させたい。

参考文献

- [1] 菅田貞治, 他: 友人関係ネットワークの変化に対応した関係構築中心性の提案, 情報処理学会研究報告, vol. 2016-MPS-107, No. 16, pp. 1-6, 2016.
- [2] 時田遼嗣, 他: ノードに内在する関係発生能力に注目したネットワーク生成モデル, 情報処理学会研究報告, vol. 2021-MPS-132, No. 9, pp. 1-6, 2021.