

## 1 はじめに

近年、社会ネットワーク分析の研究が盛んに行われているが、その中でもグラフ部分構造を詳細に分析するエゴセントリックネットワーク分析がある。社会ネットワーク分析において、ネットワーク内の少数のまとまりに着目し、ネットワーク形成の原理を見出すネットワーク・センサスというモチーフを分析する手法が用いられてきた [1]。実際に現実社会のネットワークにおいて観測されるエゴセントリックネットワークは特徴的な構造を持つパターンが観測されることが明らかになっている [2]。これらのエゴセントリックネットワークの多くは時間とともに構造を変化させていく。本稿では、時系列に伴って変化する友人関係ネットワークの局所的構造の遷移を分析する。

## 2 エゴセントリックネットワーク

各行為者を頂点とし、行為者間の影響関係を辺で表した無向グラフを考える。社会ネットワーク分析では、行為者の 1 人 1 人に注目するときに、各行為者をエゴと呼ぶ。各エゴを中心としたローカルネットワーク、つまりエゴと直接つながる行為者（オルター）の集合から誘導される部分グラフをエゴセントリックネットワーク（以下エゴネット）という。（図 1）

## 3 提案手法

同じ頂点集合（人の集合） $N$  をもつ時系列  $t_1, t_2, \dots$  に対応したグラフ系列  $G_1, G_2, \dots$  を考える。これらに観測される互いに非同型のエゴネットすべての集合を  $EN$  とする。このとき  $EN$  の与えられた分割（クラスタリング） $EN = EN_1 \cup EN_2 \cup \dots \cup EN_m$  ( $EN_i \cap EN_j = \phi$ ) に対し、対応するエゴネットの遷移図を考える。すなわち、遷移図は分割の各クラスタ  $EN_1, \dots, EN_m$  を頂点とする有向グラフであり、ある時刻  $t$  の  $G_t$  に属する人のエゴネットがクラスタ  $EN_i$  に属し、 $G_{t+1}$  でその人のエゴネットが  $GN_j$  に属するとき、 $EN_i$  から  $EN_j$  へ遷移したと見做す。有向辺は遷移の全体に対する割合に比例した確率をラベルとして持つこととする。この遷移図が本研究の対象である。

本研究では社会ネットワークの局所的推移傾向を探るためにエゴネットをクラスタリングし遷移図を得る次の 3 通りの手法を検討する。これらはそれぞれエゴネットの構造的特徴に基づくクラスタリング、ネットワークの定量的指標に基づくトップダウンな分類、

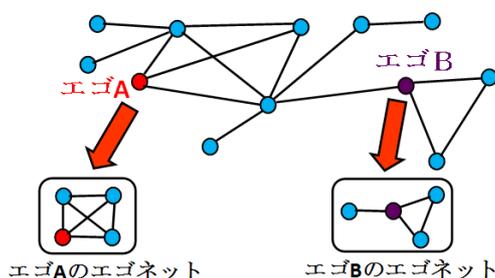


図 1: エゴセントリックネットワーク

遷移図の指標に基づく分類である。また、遷移図を定量的に評価する 2 つの指標を提案する。

### 3.1 遷移図のクラスタリングモデル

**閉路ベクトルモデル** 閉路ベクトルモデルはエゴ  $e$  のエゴネットを  $(v_1, v_2, \dots)$  に符号化する。 $v_i$  はエゴを始点かつ終点とする長さ  $i+1$  の閉路の数である。これらのベクトル間のユークリッド距離を対応したエゴ間の類似度とし、 $EN$  に階層クラスタリングを適用する。構造的特徴に基づいた距離尺度は他にもあるが、計算量の都合上、本研究では閉路による距離を用いる。

**トップダウンモデル** トップダウンモデルでは、ネットワークの定量的指標を属性値とし、各エゴネットを決定木（図 2）を用いて 8 つのクラスタに分類する。属性値の内、島はエゴネットからエゴを取り除いた際の連結成分個数、平均クラスタ係数は同値で分類されるエゴネットの平均値、グループはエゴを取り除いた際の連結成分の内、孤立ノードでないものを指す。ゲートウェイとハブは本来、コミュニティをつなげる行為者を指す単語であるが、本稿ではエゴネットに合わせており、意味が異なる [3]。

**指標優先モデル** 指標優先モデルは、後述のクラスタ指標と遷移指標の合計値（あるいは重みづけ合計）が最大の 2 クラスタ対を探索し、クラスタリングする。これをクラスタの個数が一定数になるまで繰り返す。

### 3.2 遷移図の評価指標

**クラスタ指標** クラスタ指標は、遷移図のクラスタ内で構造的に近しいエゴネットがクラスタリングされているかを表す指標である。2 グラフ間で共有する極大共通誘導部分グラフのサイズからグラフ間の類似度を定義する MCS 法 [4] に基づいた類似度のクラスタ内最速値をクラスタサイズに合わせて加重平均した値を用いる。値の範囲は  $[0, 1]$  であり、値が高いほどクラスタ内がまとまっているといえる。

**遷移指標** 遷移図のクラスタからクラスタへの各遷移がどれだけ単純化されているかを表す指標である。各クラスタの遷移確率に対するエントロピーを各クラスタサイズに合わせて加重平均し、 $[0, 1]$  の範囲で正規化する。値が高いほどクラスタ間の遷移の傾向が理解しやすいといえる。

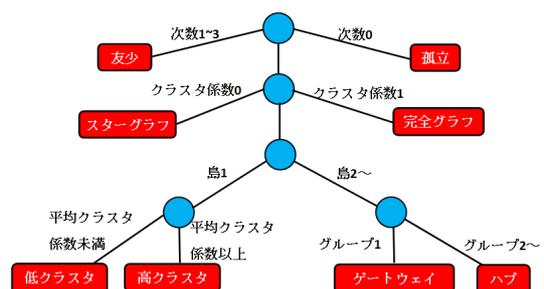


図 2: トップダウンモデルで用いる決定木

### 4 友人関係ネットワークへの適用実験

3 節で与えた手法を最終的なクラスタ数を 8 として本大学にて 4,5,6,7 月に出欠データから得た学生友人関係ネットワークより生成された遷移図に適用した。表 1 は得た遷移図の各指標値である。表内のクラスタ優先はクラスタ指標の値のみを優先して指標優先モデルを適用した場合の結果である。表内の k の値は閉路ベクトルモデルにおける符号化ベクトル長を意味する。また、クラスタ指標の%表記はクラスタ優先における値を 100%とした時の割合を表す。図 3,4 にはそれぞれトップダウンモデル、閉路ベクトルモデル (k=3) を適用した遷移図から遷移確率 12.5%以上の遷移と、エゴネットの総数に対するクラスタの割合を%表記した。閉路ベクトルモデルでは、クラスタ内の頻度 1 位、2 位のエゴネットパターンを示した。

### 5 実験結果の考察

遷移指標は指標優先モデルが最も高いが、クラスタを確認すると 1 つのクラスタにほぼ全てのエゴネットが集中して、そのクラスタ内の自己ループに遷移をかためることで遷移指標を高めているため、遷移を理解するには不適切である。クラスタ指標に関してはトップダウンモデルとクラスタ優先以外は低い値であるが、対象のネットワークを 8 個のクラスタに分けた際の取りうる最大の指標値はクラスタ優先の値になるため、トップダウンモデルは非常によくまとまっている。また、両指標はトレードオフな関係にある。遷移図から見出される傾向として、図 3 では友少クラスタとゲートウェイクラスタに含まれるエゴネットの割合が大きい。友少クラスタは社会ネットワークのスケールフリー性によるものだが、ゲートウェイクラスタは友人関係ネットワーク特有の性質と考えられる。遷移の傾向としては、クラスタ係数があまり高くないエゴネットを持つ学生は友人同士の結びつきが強固になる傾向が強い。図 4 では自己ループが高い遷移確率を持ち、また次数が近いクラスタ間に遷移があることからクラスタ係数が高いエゴネットは友人関係が安定しており、構造変化が微小であることが分かる。

表 1: モデル別指標

|          | クラスタ指標       | 遷移指標  |
|----------|--------------|-------|
| トップダウン   | 0.510(87.5%) | 0.296 |
| 閉路 (k=3) | 0.242(41.5%) | 0.535 |
| 閉路 (k=4) | 0.189(32.4%) | 0.681 |
| 閉路 (k=5) | 0.170(29.2%) | 0.741 |
| 指標優先     | 0.131(22.5%) | 0.838 |
| クラスタ優先   | 0.583(100%)  | 0.210 |

### 6 まとめと今後の課題

本研究では、友人関係の成長に関する知見を得るため、エゴネットを用いた局所的構造事変変化の分析を行った。提案したモデルの内、トップダウンモデルによるクラスタリングが有効であり、いくつかの知見が得られた。今後の課題として、評価指標の改良による

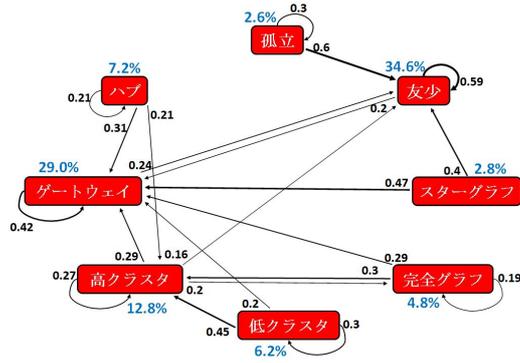


図 3: トップダウンモデル遷移図

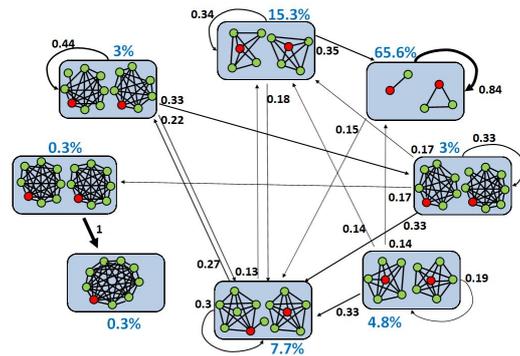


図 4: 閉路ベクトルモデル (k=3) 遷移図

有効なボトムアップ手法の提案、及び異なる友人関係ネットワークにおける遷移傾向の検証による共通傾向の発見が挙げられる。

### 参考文献

- [1] Stanley Wasserman and Katherine Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1994.
- [2] 竹内晨, 犬塚信博. エゴセントリックネットワークのパターンマイニング. 情処全大講演論文集. 2011, pp.653-655..
- [3] Bin-Hui Chou, Einoshin Suzuki. *Discovering Community-Oriented Roles of Nodes in a Social Network*. Data Warehousing and Knowledge Discovery Lecture Notes in Computer Science Volume 6263, 2010, pp.52-64.
- [4] Horst Bunke, Kim shearer. *A graph distance metric based on the maximal common subgraph*. Pattern Recognition Letters 19. 1998, pp.255-259.

### 研究業績

- 伊東樹希, 加藤竜丸, 松島裕, 武藤敦子, 犬塚信博. エゴセントリックネットワークの分析によるエゴの分類. 第 74 回情報処理学会全国大会, 2ZC-10, 2012.
- 林宏紀, 伊東樹希, 西尾典晃, 武藤敦子, 犬塚信博. エゴセントリックネットワークと形式概念分析を利用した社会ネットワーク分析. 人工知能学会論文誌 29(1), pp.177-181, 2014.
- 伊東樹希, 菅田貞治, 武藤敦子, 犬塚信博. 時間的推移を伴う友人関係の局所的ネットワーク分析. 第 76 回情報処理学会全国大会, 4L-4, 2014.(発表予定)