

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

伊東樹希[†] 加藤竜丸[†] 松島裕[‡] 武藤敦子[‡] 犬塚信博[‡]

[†]名古屋工業大学工学部情報工学科

[‡]名古屋工業大学院工学研究科情報工学専攻

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、現実の社会ネットワークは非常に規模が大きくその構造は複雑である。グラフ部分構造を詳細に分析するにはマイクロな観点からのアプローチにエゴセントリックネットワーク分析がある。

本稿では、エゴセントリックネットワークの構造的特徴に基づいたクラスタリング手法を提案する。さらに実際の友人関係ネットワークに対して提案したクラスタリング手法を適用して、ネットワークの特徴を分析する。なお、連結成分を考慮したコミュニティの抽出をネットワークのクラスタリングと言うが、本稿で取り扱うクラスタリングは類似した特徴を持つノードを分類するものであって異なる。

2 エゴセントリックネットワーク

行為者の集合を V とし、その中の各行為者 $v_1, v_2 \in V$ の間の影響関係を辺 $(v_1, v_2) \in V \times V$ で表した無向グラフ $G = (V, E) (E \subseteq V \times V)$ を考える。

社会ネットワークでは、行為者の1人1人に注目するときに、これをエゴと呼ぶ。それを中心としたローカルなネットワーク、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。即ち、 G に対する $c \in V$ のエゴセントリックネットワークとは3項組 $G_c = (c, V_c, E_c)$ であり、 $V_c = \{v \in V | (c, v) \in E \cup c\}$, $E_c = \{(u, v) \in E | u \in V_c \wedge v \in V_c\}$ である。

3 クラスタリング手法の提案

本研究では重み無し無向グラフ上のエゴをクラスタリングするため、任意の2頂点間の距離をエゴネットの構造的特徴から算出する。この算出手法には2つのアプローチを提案する。グラフ上の各エゴネットの構

造的特徴を数値化したベクトルを対応したエゴに与えそれらを基に距離を算出する方法(ベクトル手法)と、各エゴネットの特定の構造的特徴に含まれるエゴの近傍頂点群同士の距離からすべての任意の2頂点間の距離が収束するまで再帰的に算出する方法(再帰手法)である。本研究では、閉路に着目した前者の手法と、極大クリークに着目した後者の手法を考える。

3.1 閉路ベクトル手法

ベクトル手法はエゴネットの構造的特徴のみから距離を設定するため、分類結果はユーザーの視覚的観点から直感的に近いものになりやすい。閉路ベクトル手法では、エゴ e のエゴネットを (v_1, v_2, \dots) に符号化する。ここで v_i はエゴを始点かつ終点とする長さ $i+1$ の閉路の数となる。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距離とする。

3.2 極大クリーク再帰手法

各エゴ間の距離が与えられた時、別の2つのエゴ間の距離はそれらのエゴネットに含まれるオルター間の距離に依存すると考えることができる。そこで極大クリーク再帰手法では各エゴネットの極大クリーク集合に含まれるオルター間の距離から2つのエゴ間の距離を算出する。このアルゴリズムを *Algo.1* に示す。これは任意の2つのエゴ間の距離を収束するまで再帰的に算出する(6行)。この距離は極大クリーク間の距離の合計値を正規化したものであり、極大クリーク集合の要素数が異なる場合は対応しない極大クリークの距離は1である。極大クリーク間の距離は上記と同じように極大クリークに含まれるオルター間の距離の合計値を正規化した値である。

4 実験と考察

3節で与えた手法を空手クラブネットワーク(図1)に適用した。このネットワークは社会ネットワーク分析によく用いられる人間関係ネットワークである。下図は階層的クラスタリング結果を表す樹形図(デンドログラム)である。閉路ベクトル手法ではベクトルの長さは4までに限定、すなわち (v_1, \dots, v_4) とした。極大クリーク再帰手法では計算量の都合上、極大クリークの集合はサイズが大きいものから順に最大で4つま

A Classification of the ego by analysis of an egocentric network

Tatsuki Itou[†], Ryumaru Kato[†], Hiroshi Matsushima[‡], Nobuhiro Inuduka[‡] and Atsuko Muto[‡]

[†]Dept. of Computer Science, Faculty of Engineering, Nagoya Institute of Technology

[‡]Dept. of Computer Science and Engineering, Graduate School of Engineering

Algorithm1 Maximal clique recurrence

入力: $G = (V, E)$

出力: $dist(e_1, e_2), e_1, e_2 \in V$

1. 任意の $e_1, e_2 \in V$ で距離 $dist(e_1, e_2)$ を初期化
 2. Repeat
 3. For each $(e_1, e_2) \in V \times V$ do.
 4. $S_1 = e_1$ の G におけるエゴネットのすべての極大クリークの集合
 $S_2 = e_2$ の G におけるエゴネットのすべての極大クリークの集合
 5. if $|S_1| < |S_2|$ then S_1 と S_2 を交換
 6. $dist(e_1, e_2)$ を任意の単射 $S_2 \rightarrow S_1$ での次の値の最小値に更新

$$\frac{1}{|S_1|} (\{ \sum_{K \in S_2} dist(f(K), K) \} + (|S_1| - |S_2|))$$
 7. end for
 8. Until どの (e_1, e_2) にも更新がない
- ただし、
 $dist(K_1, K_2)$ (K_1, K_2 はクリーク、 $|K_1| \geq |K_2|$) は、
 任意の単射 $g: K_2 \rightarrow K_1$ での次の値の最小値

$$\frac{1}{|K_1|} (\{ \sum_{e \in K_2} dist(g(e), e) \} + (|K_1| - |K_2|))$$

で限定した。なお、どの階層的クラスタリング手法も最遠離法を用いている。

図2, 図3より、前者の手法において近い距離でクラスタを形成しているエゴ群が、後者の手法では大きく分かれていることが分かる。これは各エゴネットの構造的特徴が他のエゴ間の距離に大きく影響を及ぼしているためである。そのため、同型のエゴネットでも後者では含んでいるオルターに応じてクラスタリング結果が大きく異なっている。

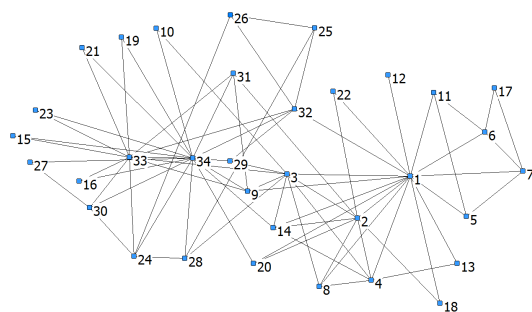


図 1: 空手クラブネットワーク

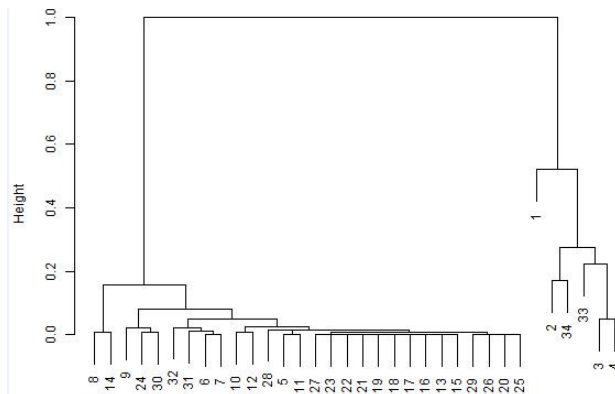


図 2: 閉路ベクトル手法による分類結果

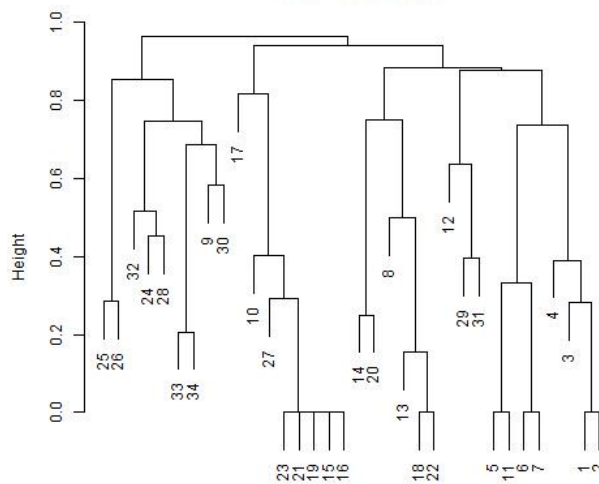


図 3: 極大クリーク再帰手法による分類結果

5 おわりに

本稿では、エゴネットの構造的特徴の観点からネットワーク上のエゴのクラスタリング手法を提案した。これを友人関係ネットワークに適用することで、社会的関係に重みをおいた社会ネットワーク分析を行った。

参考文献

[1] 神鷲敏弘 「データマイニング分野のクラスタリング手法」人工知能学会誌 18 巻 1 号, pp.59-65.
 [2] 竹内農 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」全国大会講演論文集 2011, pp.653-655.