F-056

ネットワーク間狭小性によるネットワークの時系列解析

森岡　淳†　　古川　正志‡　　山本　雅人‡　　鈴木　育男‡
†‡北海道大学大学院情報科学研究科

1 はじめに

人々は様々な人と多様な形で関わり合い，社会を形成する．その関係をネットワークとして捉え，その構造の特性から人々のコミュニケーションを理解するための様々な研究が行われている．Watts ら [1] はそのようなネットワークから，スモールワールド性を発見した．

また，関係性が絶えず変化するネットワークの解析には，時間変化を考慮する必要があり，複数のアプローチから解析が行われている．近年では Web などによりデータの時間遷移による変化の追跡が容易になったため，それらのネットワークを時系列解析しユーザーの振る舞いの変化などを観測する研究が盛んに行われている [2].

本研究では，時間遷移によって各ノードが結合しているノードがどのように変化しているかを調べ，ノードの振る舞いの変化を調べる．そのために実データから時系列毎にネットワークを作成し，二つのネットワークのノード間関係の類似度を定量化するネットワーク間狭小性 [3] をノード毎に算出して類似度を測定し，解析を行う．

2 ネットワーク間狭小性 [3]

2.1 同一ノードを持つネットワーク間の定義

本研究で使用する2つのネットワークを，G_i，G_j とする．ここで，V_i は G_i の頂点集合，E_i は G_i の辺集合である．ネットワーク $G_i = (V_i, E_i)$ のノードが持つラベルを $\psi_i(x)$ と書く．全ての $x \in V_i$，$y \in V_i$ について，$\psi_i(x) = \psi_j(y) \Leftrightarrow x = y$ とする．任意の2ネットワーク G_i，G_j について $\forall x \in V_i$，$\forall y \in V_j$ であるならば，$\psi_i(x) = \psi_j(y) \Leftrightarrow x \sim y$ とする．以降，このような x，y を同一ノードとする．

2.2 ネットワーク間狭小性の定量化

まず $v(v \in G_i, G_j)$ の G_i，G_j におけるパス長 k 内の近傍，$\Gamma_i^k(v)$，$\Gamma_j^k(v)$ の内，同一ノードが含まれる割合 NS_{ij}^k を式 (1) で定義する．

$$NS_{ij}^k(v) = \frac{\left|\Gamma_i^k(v) \cap \Gamma_j^k(v)\right|}{\min\left(\left|\Gamma_i^k(v)\right|, \left|\Gamma_j^k(v)\right|\right)} \frac{1}{k} \tag{1}$$

$\Gamma_i^k(v)$ は v_l のネットワーク G_i での k 次近傍内のノードの集合である．$\min\left(\left|\Gamma_i^k(v)\right|, \left|\Gamma_j^k(v)\right|\right)$ は両ネットワークでのパス長 k 内の近傍ノード数のうち小さいものであり，$\frac{1}{k}$ は近傍を広げるにつれて，ネットワーク狭小性の増分を抑える作用がある．

NS_{ij}^k を用いて，ネットワーク間狭小性を定量化する．ネットワーク G_i，ネットワーク G_j 間の同一ノード v のネットワーク間狭小性 $NS_{ij}(v)$ を式 (2) に示す．

$$NS_{ij}(v) = \frac{1}{\sum_{k=1}^{R} k^{-1}} \sum_{k=1}^{R} NS_{ij}^k(v) \tag{2}$$

ここで $R = \max\{\mathrm{rad}(v|i), \mathrm{rad}(v|j)\}$ であり $\mathrm{rad}(v|i)$ は i におけるノード v の半径である．$k = R$ のときは，v が G_i，G_j で到達できる全てのノードが $\Gamma_i^R(v)$，$\Gamma_j^R(v)$ に含まれる．$\sum_{k=1}^{R} k^{-1}$ は，$NS_{ij}(v_l)$ の最大値を1とするように正規化するためである．これを両ネットワークに含まれる全ての同一ノードについて計算しネットワーク G_i，ネットワーク G_j 全体のネットワーク間狭小性 NS_{ij} を式 (3) で求める．

$$NS_{ij} = \frac{1}{n_{ij}} \sum_{l=1}^{n_{ij}} NS_{ij}(v_l) \tag{3}$$

式中の n_{ij} はネットワーク G_i，ネットワーク G_j にともに存在するノードの数である．式 (3) は，式 (2) の全同一ノード平均の算出に相当する．以後，この値をネットワーク間狭小性とし，NS と略す．

3 時系列解析

3.1 ネットワーク作成に用いたデータ

本研究で用いたデータは，ソーシャルブックマークサービスの del.icio.us(http://del.icio.us) から取得した．期間は 2007 年 1 月から同年 12 月までの間である．取得したユーザー数は 1,000，エントリ数は 1,763,060 となった．ここから，期間を3ヶ月ごと4つに分け，ユーザーをノードとし，同じエントリをブックマークしたユーザー同士にリンクを張り，4つのネットワークを作成した．表 1 に各ネットワークのリンク数を示す．

Time Series Analysis of Network Based on Network Small-worldness Quantification
†Jun Morioka　‡Masashi Furukawa　‡Masahito Yamamoto　‡Ikuo Suzuki
†‡Graduate School of Infomation Science And Technology, Hokkaido University

表 1: 各ネットワークのネットワークのリンク数

対象期間	リンク数
1月〜3月	87,829
4月〜6月	91,264
7月〜9月	72,542
10月〜12月	59,146

表 2: 各組のネットワーク間狭小性

ネットワークの組	狭小性
1月〜3月：4月〜6月	0.77241
4月〜6月：7月〜9月	0.77499
7月〜9月：10月〜12月	0.78754

3.2 ネットワーク間狭小性の算出

4つのネットワークから3つのネットワークの組み合わせを選び，ネットワーク間狭小性を算出した．3つの組み合わせは，時間順に選んだ．つまり，1月から3月のネットワークと4月から6月のネットワーク，4月から6月と7月から9月のネットワーク，7月から9月と10月から12月のネットワークの3組である．各組のネットワーク全体の狭小性を表2に示す．どの組も高い値を持ち，あまり差が無いことがわかる．

次に，同じ3組からノード毎のネットワーク間狭小性を算出した．結果を図1に示す．縦軸は狭小性，横軸は3組での狭小性の標準偏差が低いものから順位付けしたものである．なお，図を見やすくするために順位が10の倍数のものだけを表示した．また，全ての組で狭小性が0のものは除外してある．結果を見てみると，順位の高いものほど，狭小性も高いことがわかる．これより，狭小性の高いユーザーほど，常に同じユーザーと関係を持ちやすいことがわかる．また，順位が600前後から大きく値が下がっているが，これはいずれかの組で狭小性が0になっているノードである．

さらに，ネットワークの組の推移，つまり時間の推移により狭小性がどのように変化しているかを調べる．そのために，順位を200ごとに区切り，各々から5つのノードを無作為に選び図2に示した．縦軸は狭小性，横軸はネットワークの組である．この図より，最初は狭小性が高いが最終的に低いユーザー，最初と最後は狭小性が低いが中間だけ高いユーザーなど，様々な特徴が見受けられる．興味深いのは狭小性が常に低いノードがないこと，つまり常に違うユーザーと関係を持っているようなユーザーがいないという点である．

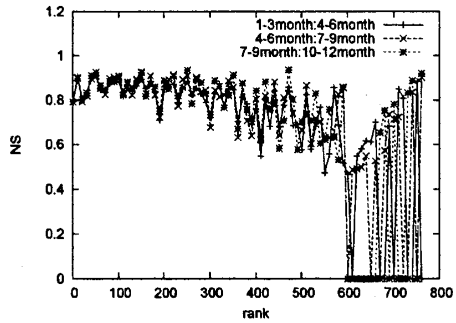

図 1: ノードのネットワーク間狭小性

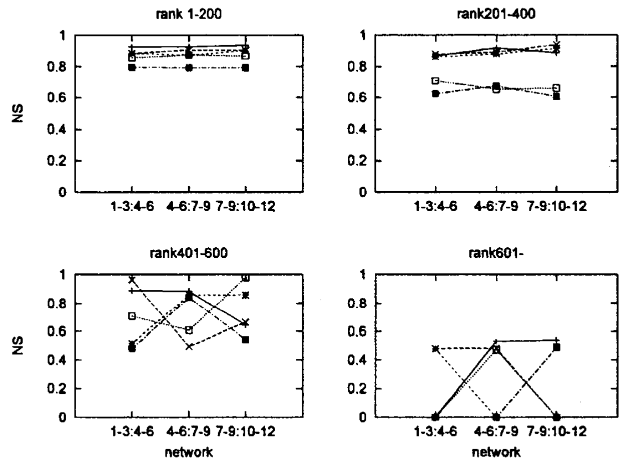

図 2: ノードのネットワーク間狭小性の推移

4 おわりに

本研究では，時系列毎のネットワークを作成し，それらからノード毎のネットワーク間狭小性を算出することで，各ユーザーの振る舞いが時間遷移によりどのように変化しているかを調べることが可能であることを示した．今後は他の特徴量と照らし合わせ，より詳細に解析していくことが重要である．

参考文献

[1] D. J. Watts, Small Worlds: The Dynamics of Networks Between Order and Randomness, Princeton Univ. Press, Princeton, NJ, (1999).

[2] 河内佑美，吉井伸一郎，ビヘイビアダイナミクスの複雑ネットワーク的解析，日本ロボット学会誌，vol.26，No.1，19-22(2008)

[3] 森岡淳，古川正志，山本雅人，鈴木育男ノード間関係の類似度を定量化するネットワーク間狭小性，情報処理学会第70回全国大会，2，27-28