2B4-NFC-02b-2

ネットワーク構造が情報拡散に与える影響の分析

Influence analysis of the diffusion networks structure

臼井 翔平[*1*2] Shohei Usui　鳥海 不二夫[*1] Fujio TORIUMI　松尾 真人[*3] Masato MATSUO　平山 高嗣[*2] Takatsugu HIRAYAMA　間瀬　健二[*2] Kenji MASE

[*1] 東京大学大学院工学系研究科 School of Engineering The University of Tokyo

[*2] 名古屋大学大学院情報科学研究科 Graduate School of Information Science, Nagoya University

[*3]NTT 未来ねっと研究所 NTT Network Innovation Laboratories

Successful diffusion of information on Social Media draw the attention after the 2011 off the Pacific coast of Tohoku Earthquake at 2011. In order to use the Social Media as information sources safely and swift, we need to analyze diffusion of information. In this paper, we focus on network topologys which we undestand how infomation diffuse. We analyze how infomation diffusion depends on the network topologys. We propose a network growth model that can generate networks of the expected features. As a result, we confirm that the number of users who get infomation is affected by value of assortativity and degree distribution.

1. はじめに

近年，ソーシャルメディア上での情報拡散が社会に与える影響が注目されている．2011 年 3 月 11 日に発生した東日本大震災の際は，Twitter が情報を伝達，収集するツールとして機能したことは記憶に新しい．そのような中，災害時の新たな情報源としてのソーシャルメディアが期待を集めている．そのために，現在ソーシャルメディア上での情報拡散の分析が行われている．

本研究では，情報拡散が行われるネットワークの構造に着目し分析を行う．鳥海らは震災前後のユーザのコミュニケーションからネットワークを抽出し，その構造が震災の前後で変化している事を示している [1]．このような構造の変化は情報拡散のしやすさに影響を及ぼすと考えられる．そこで本研究では，ネットワーク構造と情報拡散のしやすさの関係についての分析を行う．

まず，ネットワーク構造を既存のネットワーク指標を用いて定量的に表現する．次に，任意に与えた指標値 (以下，目的指標値) に近い値を持つネットワークを生成できるネットワーク成長モデルを提案する．各々の指標が情報拡散に与える影響を明らかにするため，提案したモデルを用いてそれぞれを変化させたネットワークを生成する．提案モデルによって生成される，様々な構造を持つネットワーク上で情報拡散シミュレーションを行い，構造と情報拡散の間の関係を分析する．

2. ネットワーク成長モデルの提案

2.1 概要

本研究では，目的指標値に近い指標値を持つネットワークを生成するネットワーク成長モデルを提案する．提案モデルでは，**図 1** のように，n ステップ目のネットワークに対し，複数のアルゴリズムによって $(n+1)$ ステップ目のネットワーク候補を生成する．次に，全ての $(n+1)$ ステップ目のネットワーク候補の中で，評価関数によって最適なネットワークを選択し，$(n+1)$ ステップ目のネットワークとする．このステップを繰り返すことによって，目的指標値に近い指標値を持つネットワークを生成する．

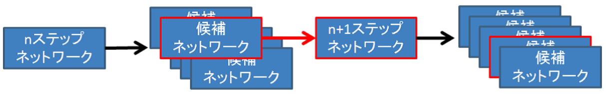


図 1: ネットワークの成長

2.2 ネットワーク指標

本研究では，ネットワーク構造を定量的に表現するために以下に示す既存の指標の値をネットワークから算出する．

(1) 平均経路長 L [2]

(2) クラスタ係数 C [2]

(3) 同類選択性 r [3]

(4) べき指数 γ [4]

2.3 評価関数

$(n+1)$ ステップ目のネットワーク候補の中から最適なネットワークを選択するためには，目的指標値と生成中のネットワークの指標値の差を定量的に評価する必要がある．本研究では，使用する全ての指標値から類似距離 D を算出する [8]．なお，この際ノードの次数分布がどれだけべき分布に近いかを表す，べき分布決定係数 R^2 も同時に用いる．

$$D(G_i) = (\frac{L_i - L_a}{\sigma_L})^2 + (\frac{C_i - C_a}{\sigma_C})^2 + (\frac{r_i - r_a}{\sigma_r})^2 + (\frac{\gamma_i - \gamma_a}{\sigma_\gamma})^2 + (\frac{R_i^2 - {R^2}_a}{\sigma_{R^2}})^2 \quad (1)$$

ただし，L_i，C_i，r_i，γ_i，${R^2}_i$ は，それぞれネットワーク G_i の指標値であり，L_a，C_a，r_a，γ_a，${R^2}_a$ は，それぞれ目的指標値である．各項の分母 $\sigma_L, \sigma_C, \sigma_r, \sigma_\gamma, \sigma_{R^2}$ は，各指標の標準偏差を表す．これは各指標を正規化するための係数である．

連絡先: 名古屋大学大学院情報科学研究科
〒 464-8603 愛知県名古屋市千種区不老町
E-mail: usui@cmc.ss.is.nagoya-u.ac.jp

2.4 ネットワークの成長

提案モデルでは，各ステップでノードの追加かリンクの追加のいずれかの操作を選択し，次ステップ目の候補ネットワークを生成する．これらの操作の中でリンクを作成する際，既存ノードの中からノードを選択する．その際，3 つの要素を組み合わせたアルゴリズムにより複数の候補ネットワークを生成する．以下では，まず組み合わせ要素について説明し，次にノードの追加及びリンクの追加の際の操作について説明する．

2.4.1 ノード選択アルゴリズム

提案モデルでは以下の 3 つの要素を用いてノードを選択する．

- 次数優先性 q_k[-1,0,1]
- 適応度優先性 q_η[0,1]
- 近傍優先性 [true,false]

ノード i が選択される確率 Π_i は，ノード i の「次数 k_i」と「適応度 η_i」を用いて式 (2) で表される．

$$\Pi_i = \frac{k_i^{q_k}\eta_i^{q_\eta}}{\sum_j k_j^{q_k}\eta_j^{q_\eta}} \tag{2}$$

なお，適応度は各ノードに一様分布になるように設定されたパラメータである．また，$q_n = true$ の際は，1 つのノードが選択されている際にそのノードと経路長が 2 であるノード群の中から選択され，$q_n = false$ の際は全ての既存ノードの中から選択される．

次数優先性

実社会の多くのネットワークは強いスケールフリー性を持っている．このようなネットワークが生成される過程では，多くのリンクを獲得しているノードはリンクを獲得しやすいという性質が考えられる [4]．一方，リンクの少ないノードはリンクを獲得しようとする．このような性質を次数優先性によって表現する．

適応度優先性

提案モデルでは，ノード毎に「他のノードと次数を競う能力」を表す「適応度（fitness）η」を設定する [5]．この適応度が高い程次数を獲得しやすくなる性質を適応度優先性によって表現する．

近傍優先性

近年の SNS には友人の推薦システムにより，友人の友人が推薦される事が多い．また，推薦システム以外にも友人の友人を SNS 上で見かける事は多く，「友人の友人は，友人になりやすい」と考えられる [6]．このような性質を近傍優先性によって表現する．

2.4.2 ノードの追加

ネットワークに新たなノードを 1 つ追加し，既存のノードとの間にリンクを形成する．その際，以下の要素を考える．

- 次数優先性
- 適応度優先性

従ってノードの追加では，次数優先性 3 通りと適応度優先性 2 通りを組み合わせた 6 通りのノードの選択手法を用いて候補ネットワークを生成する．この際，例えば次数優先性が正で適応度優先性が $true$ である場合．Π_i は次数と適応度の双方に比例する．

2.4.3 リンクの追加

既存のノードを 2 つ選択し，それらのノード間にリンクを形成する．まず，以下の要素を用いて，1 つ目のノードを選択する．

- 次数優先性
- 適応度優先性

従って，1 つ目のノードは次数優先性 3 通りと適応度優先性 2 通りを組み合わせた 6 通りのノードの選択手法を用いて選択する．

次に，2 つ目のノードを上記の要素に加え以下の要素を考えて選択する．

- 近傍優先性

従って 2 つ目のノードは，次数優先性 3 通り，適応度優先性 2 通り，近傍優先性 2 通りを組み合わせた 12 通りのノードの選択手法を用いて選択する．

2.4.4 ネットワーク生成手順

提案モデルでは以下の手順でネットワークの生成を行う．

(1) n ステップ目のネットワークに対してノードの追加かリンクの追加のいずれかを選択し，ネットワーク候補を生成
   (a) ノードの追加 (確率 P)
   (b) リンクの追加 (確率 $1 - P$)

(2) 各アルゴリズムによって作られたネットワーク候補に対して類似距離 D を計算

(3) D が最も小さいネットワークを $(n+1)$ ステップ目のネットワークとする

この手順を繰り返し行うことでネットワークを生成する．
リンク数が求めるネットワークのリンク数に達した時点で試行を停止する．なお，確率 P は生成するネットワークのノード数とリンク数から決定される．

3. 情報拡散シミュレーション

3.1 シミュレーション概要

本章では，第 2.2 節で挙げた指標が変化することで，情報拡散がどのように変化するかを分析する．そのために，第 3 章で提案したネットワーク成長モデルを用いて，各指標を変更しながらネットワークを生成し，生成されたネットワーク上で，情報拡散モデルを用いて情報拡散シミュレーションを行う．ここで，情報を受け取ったノードの全体に占める割合を期待影響度とし，情報拡散が収束した後の期待影響度を終端期待影響度とする．終端期待影響度が大きい程，当該ネットワーク上では情報が広く拡散されていることを示している．

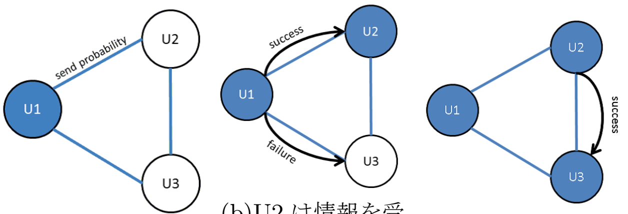
3.2 シミュレーション設定

まず，各指標値を対象とし，その指標のみが情報拡散に与える影響を分析するため，1 つの初期ネットワークから得られる指標値を基準とし，対象指標値のみを変化させたネットワークを生成する．ここで，初期ネットワークとして「So-net SNS」(β 版)*1に存在していた友人ネットワークを利用した．なお，平均経路長は 1.0 ずつ，それ以外の指標値は 0.1 ずつ，**表 1** に

*1 http://so-netsns.jp/

表 1: 指標の設定

	初期ネットワーク	範囲
平均経路長 L	3.59	$1.0 \leq L \leq 10.0$
クラスタ係数 C	0.227	$0 \leq C \leq 1.0$
同類選択性 r	-0.227	$-1.0 \leq r \leq 1.0$
べき指数 γ	1.31	$0 \leq \gamma \leq 2.0$
べき分布決定係数 R^2	0.889	-



(a)U1 は始点ノード　(b)U2 は情報を受け取ったが U3 は受け取らない　(c)U2 から U3 に情報が伝わる

図 2: AsIC モデルの操作

示す範囲で変化させ，類似距離 D が 1 未満であるネットワークを生成できた範囲のみを使用した．また，統計値を出すために 1 つの指標値セットに対して 100 回ネットワークを構築しシミュレーションを行う．

情報拡散シミュレーションには吉川らが提案した AsIC モデルを使用する [7]．なお，シミュレーション開始時に情報を所持しているノードを始点ノードと呼ぶ．AsIC モデルは確率モデルであるため，各ノードを始点ノードとして，100 回ずつ，全てのノードに対して試行し終端期待影響度の平均を算出する．

3.3 AsIC モデル

AsIC モデルは，感染病の広がり方を示す基本的な確率モデルである IC モデルを，非同期時間遅れを考慮するように拡張したモデルである [7]．**図 2** に AsIC モデルの基本的な流れを示す．U1 が始点ノードのとき，隣接ノードである U2 と U3 に対し 1 度だけ情報を発信する．この際 U1 はノード間のリンクに設定された発信確率に従って情報を発信するか否かを決定する．ユーザに情報が伝わると，遅延時間が経過した後に，受信ユーザが発信源となり，再び 1 度だけ情報発信を行う．この操作を繰り返すことにより情報拡散を表現する．

3.4 シミュレーション結果

3.4.1 相関係数

生成した全てのネットワークに対して，変化させている指標と，終端期待影響度の変化の間の相関係数をそれぞれ**表 2** に示す．**表 2** より，平均経路長 L が最も強い負の相関を示している事がわかる．平均経路長はネットワークの大きさを表す指標であり，この値が大きくなればノード間の距離が大きくなるため，情報が伝わりにくくなるのは明らかである．また，同類選択性 r やべき指数 γ に強い正の相関が得られ，クラスタ係数 C は弱い正の相関を示している．これら 3 つの指標について以下で細かく分析する．分析に際し，同一目的指標値毎に終端期待影響度の平均を算出する．

3.4.2 同類選択性 r

同類選択性 r を変化させながら生成したネットワーク上での終端期待影響度を**図 3** に示す．r が 0 に近づく程，終端期待影響度は大きくなることがわかる．r が小さい時は次数の小さいノードと大きいノードがリンクする傾向にあり，r が大きい時は次数の大きいノード同士，次数の小さいノード同士がリンクする傾向にある．従って，r が小さい時は多くのノードが次数の大きいハブノードと結びついているのに対し，r が大きい時はハブノードにリンクしているノード同士がリンクしている事が示唆される．**図 4**(a) に r が小さい時の，**図 4**(b) に r が大きい時のネットワーク例を示す．これらはノードを次数順にソートし円形に並べたネットワーク図である．ここからも，**図 4**(b) ではハブノード以外のノード同士が多く結びついている事が言える．以上より，ハブノード以外のノード同士が結びついているネットワークで情報拡散がされやすいことが示唆される．これは，ハブノードの周りのノードが結びつくことによって，情報を受け取るチャンスが増加するためであると考えられる．

表 2: 各指標値と終端期待影響度の相関

L	C	r	γ
-0.839	0.340	0.627	0.696

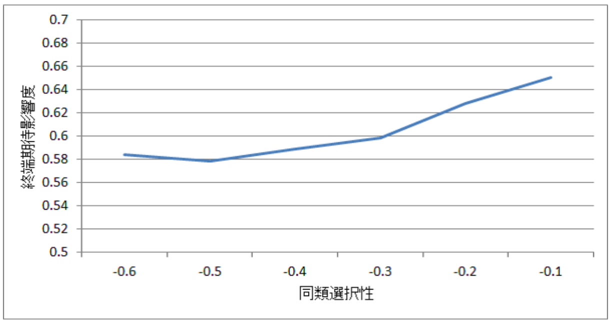
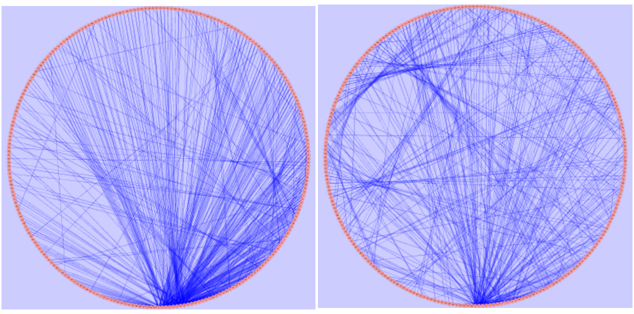


図 3: 終端期待影響度：同類選択性

3.4.3 べき指数 γ

べき指数 γ を変化させながら生成したネットワーク上での終端期待影響度を**図 5** に示す．γ が高くなるにつれて，終端期待影響度は大きくなることがわかる．**表 3** に次数が上位 3 位以内のノードの累積次数割合の平均を示す．べき指数が小さいほど上位ノードへの次数の偏りが大きいと言える．従って，γ が小さい時は多くのリンクが少数のノードに集中しているのに対し，γ が大きい時は複数のノードに次数が分散している事が示唆される．**図 6**(a) に γ が小さい時の，**図 6**(b) に γ が大きい時のネットワーク例を示す．ここからも，**図 4**(a) では多くのリンクが少数のハブノードに集中し，大きなハブノードが存在している．一方**図 4**(b) では，1 つのハブノードの大きさは小さいが，複数のハブノードが存在していると言える．以上よ



(a)r が小さい時　(b)r が大きい時

図 4: ネットワーク例

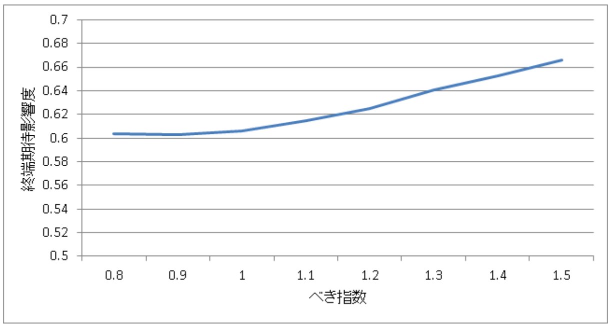


図 5: 終端期待影響度：べき指数

表 3: 次数割合最上位 3 ノード：べき指数

べき指数	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5
累積次数割合	0.184	0.197	0.196	0.183	0.168	0.162	0.137	0.119

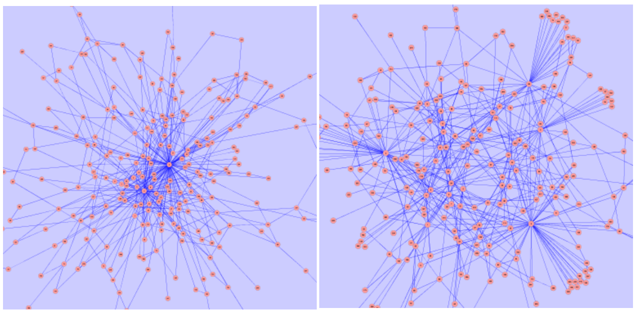
り，少数の大きなハブノードが存在するネットワークよりも，複数のハブノードに次数が分散しているネットワークで情報拡散がされやすいことが示唆される．これは，ハブノードが多くあれば，ハブノードに情報が伝わる確率が高くなるためであると考えられる．

3.4.4　クラスタ係数 C

クラスタ係数 C を変化させながら生成したネットワーク上での終端期待影響度を**図 7** に示す．r や γ に比べると変化量が小さいが，C が大きくなるにつれて終端期待影響度は大きくなることがわかる．C が大きい時，ネットワークの密度が高くなる．したがって，ネットワークの密度が高い程，情報拡散は行われやすいが，ネットワークの密度は，情報拡散に大きく影響することはない事が示唆された．

4.　結論

本研究では，ネットワーク構造の違いに着目し，情報拡散のしやすさを分析した．そのために，目的指標値を持つネットワークを生成するネットワーク成長モデルを提案した．提案したモデルを用い，各指標値を対象とし，対象指標値のみを変化させたネットワークを生成した．そして，生成されたネットワーク上で AsIC モデルを用いて情報拡散シミュレーションを行った．その結果，平均経路長に最も強い負の相関を示し，同類選択性やべき指数にも強い正の相関を示した．これらの分析より，ハブノード以外のノード同士が結びついているネットワークで情報拡散がされやすいことが示唆された．また，複数のハブノードに次数が分散しているネットワークで情報拡散がされやすいことが示唆された．



(a)γ が小さい時　(b)γ が大きい時

図 6: ネットワーク例

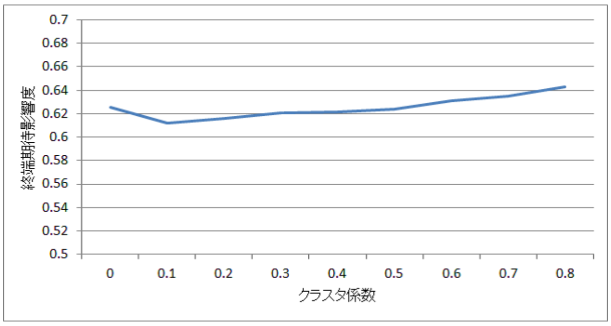


図 7: 終端期待影響度：クラスタ係数

今後の課題としては，情報拡散が起こりやすいネットワークを生成することを挙げる．また，本研究では方向のないリンクを用いた無向ネットワークを用いた．しかし，Twitter のようなソーシャルメディアにおける情報拡散には方向が存在する．そのため，方向のあるリンクを用いた有向ネットワークでの分析が求められる．

謝辞

本研究は NTT との共同研究，および科研費（24300064）の助成を受けた．また，本研究を行うにあたって「So-net SNS」(β 版) のデータを御提供頂いたソネットエンタテインメント株式会社に感謝致します．

参考文献

[1] 鳥海不二夫, 篠田孝祐, 栗原聡, 榊剛史, 風間一洋, 野田五十樹. 震災がもたらしたソーシャルメディアの変化.JWEIN11,pp.41-46,2011.

[2] DJWatts and SH Strogatz. Collective dynamics of 'small-world' networks. Nature,Vol. 393, No. 6684, pp. 409—10, 1998.

[3] MEJ Newman. Mixing patterns in networks. Physical Review E, Vol. 67, No. 2, p. 26126, 2003.

[4] A.L. Barabäsi and R. Albert. Emergence of Scaling in Random Networks. Science,Vol. 286, No. 5439, p. 509, 1999.

[5] G. Bianconi and A.L. Barabäsi. Competition and multiscaling in evolving networks. Europhysics Letters, Vol. 54, No. 4, pp. 436—442, 2001.

[6] J. Davidsen, H. Ebel, and S. Bornholdt. Emergence of a Small World from Local Interactions: Modeling Acquaintance Networks. Physical Review Letters, Vol. 88,No. 12, p. 128701, 2002.

[7] 吉川友也, 斉藤和巳, 元田浩, 大原剛三, 木村昌弘. 情報拡散モデルに基づくソーシャルネットワーク上でのノードの期待影響度曲線推定法. 信学会 (D),Vol. J94—D No. 11, pp. 1899—1908.

[8] 鳥海不二夫, 神谷達幸, 石井健一郎. 小規模 SNS におけるネットワーク成長パラメータの推定. 電気情報通信学会 2011