

コミュニケーション活動を考慮した SNS モデルの提案 Proposal of SNS Model Considering User's Communication

石田 健†
Ken Ishida

鳥海 不二夫†
Fujio Toriumi

石井 健一郎†
Kenichiro Ishii

1. はじめに

近年, mixi, Gree といった SNS (Social Networking Service) の拡大は社会現象となっており, 社会に大きな影響を及ぼしている. また, SNS の拡大に伴い, SNS 上の現象を分析するための研究や, ユーザに対する調査などが行われている [1]. しかし, 実際のユーザを対象とした研究や調査は, 時間や費用の面で制約が大きい.

そこで, 本研究では SNS を仮想的に表現するマルチエージェントモデルを提案し, それを用いたシミュレーションによって SNS 上の様々な現象を分析する. 提案モデルでは, SNS をエージェント, 及びエージェント間のリンクからなるネットワークとして表現する. さらに, ユーザ間のコミュニケーション活動をモデル化することで, 現実の SNS に忠実なモデルを実現する.

さらに, 本研究では, 提案モデルと, 名古屋大学大学院情報科学研究科において実際に運用されている学内 SNS である, Academic Community System (ACS) [2] とを比較し, モデルの妥当性を検証する.

2. 提案モデル

2.1 概要

本研究ではマルチエージェントモデルを用いて SNS 内の個々のユーザの動作を記述し, それによってネットワーク全体の振る舞いを表現する. 本モデルで考える SNS は以下の機能を持つものとする.

- SNS 内でコミュニケーション活動 (日記更新, 及びそれに対するコメント) を行う機能.
- 他のユーザを友人とし, リンクを作成する機能.

なお, 提案モデルでは, エージェントが日記やコメントの読み書きなど, 一連の動作を行う単位時間を 1 ターンと定義し, これを現実世界の 1 日とする.

2.2 エージェントモデル

SNS ネットワーク上に m 人のエージェントが存在するとき, それぞれのエージェントを $a_i (i = 1, 2, \dots, m)$ とする. エージェント a_i は以下の 3 つの内部パラメータを持つ.

- ログイン頻度 $l_i (0 \leq l_i \leq 1)$
- 日記更新頻度 $d_i (0 \leq d_i \leq 1)$
- 特徴ベクトル $\mathbf{v}^i = [v_1^i, v_2^i, \dots, v_n^i]$

ログイン頻度 l_i は, あるターンにおいてエージェントが SNS にログインする確率を表す. 例えば, ログイン頻度 $l_i = 1.0$ であれば毎ターンログインすることを表し,

$l_i = 0.0$ であれば全くログインしないことを表す. ログイン頻度の値は, 平均 μ_l , 標準偏差 σ_l の正規乱数で初期化する. すなわち, μ_l, σ_l はログイン頻度初期化用の定数である.

日記更新頻度 d_i は, エージェントがログインした際に日記を更新する確率を表す. 日記更新頻度 $d_i = 1.0$ であれば, エージェント a_i はログインした際には必ず日記を更新することを表し, $d_i = 0.0$ であれば, エージェント a_i はログインしても全く日記を更新しないことを表す. 日記更新頻度の値は平均 μ_d , 標準偏差 σ_d の正規乱数で初期化する. すなわち, μ_d, σ_d は日記更新頻度初期化用の定数である.

特徴ベクトル \mathbf{v}^i はエージェント a_i の個性を表し, 各要素は 0 または 1 の値をとる. 各要素は, ベクトルの全要素から B 個の要素をランダムに選択し, 選択された要素を 1, それ以外の要素を 0 とすることで初期化する. ここで, B は特徴ベクトル初期化用の定数である.

2.3 リンクモデル

あるエージェント a_i とエージェント $a_j (j \neq i)$ とが友人関係にある場合, その関係を無向リンク L_{ij} として表現する. 無向リンク L_{ij} はユーザ間の繋がりの強さを表す好感度 f_{ij} をもつ. 提案モデルでは, 好感度 f_{ij} を次式で定義する.

$$f_{ij} = \alpha \log(c_{ij} + 1) + \beta \quad (1)$$

(ただし f_{ij} の上限は 1)

ここで, α は好感度変化率, β は好感度の下限値を表す定数である. また, c_{ij} はエージェント a_i と a_j の間で直近の T ターンにコメントの書き込み, すなわちコミュニケーションが行われた回数である. これにより, エージェント間で頻繁にコミュニケーションが行われているほど好感度が高くなることを表現する.

図 1 に $\alpha = 0.9, \beta = 0.1, T = 14$ とした時の, コミュニケーション回数と好感度の関係を示す.

2.4 ネットワーク構築モデル

提案モデルは CNN モデル [3] と Fitness モデル [4] を組み合わせ, 「新規エージェントの追加」及び「新規リンクの作成」を確率的に繰り返すことによってネットワークを構築する.

新規エージェントの追加

既存のネットワーク内から選択されたエージェントと, 新規エージェントとの間にリンクを作成する. さらに, 選択された既存エージェントの隣接エージェントと, 新規エージェントとの間に, 「潜在的リンク」を作成する. ここで, 「潜在的リンク」は今後友達になる可能性のある関係を表す.

新規エージェントにリンクで繋がれるエージェントは, ルーレット選択によって決定される. エージェント a_i が

†名古屋大学大学院情報科学研究科

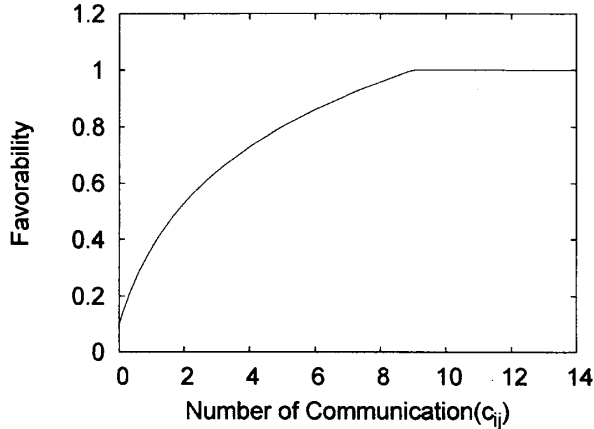


図 1: コミュニケーション回数と好感度の関係

新規エージェントとリンクで繋がれる確率 Π_i は, SNS を活発に利用しているユーザほど, 新規ユーザを SNS に招待する可能性が高くなることを表現するため, 以下の式で定義する.

$$\Pi_i = \frac{U_i k_i}{\sum_j U_j k_j} \quad (2)$$

ここで k_i はエージェント a_i の次数, U_i はエージェント a_i の SNS 活用度である. SNS 活用度 U_i はエージェントが SNS をどれだけ活用しているかを表す値であり, 以下の式で定義する.

$$U_i = \frac{\sum_{n=1}^H h_i(n)}{H} \quad (3)$$

ここで, $h_i(n)$ は n ターン前のコミュニケーション活動 (日記またはコメント) の履歴を示す値であり, 以下の値をとる.

$$h_i(n) = \begin{cases} 1 & (n \text{ ターン前に活動した}) \\ 0 & (n \text{ ターン前に活動していない}) \end{cases} \quad (4)$$

また, H はエージェントの活動履歴を何ターンさかのぼって調べるかを表す定数である.

新規リンクの作成

ネットワーク内の「潜在的リンク」を1つ選択し, それを実際のリンクに変更する. 選択される「潜在的リンク」は, 「潜在的リンク」で接続されている2エージェント間の類似度 S_{ij} を適応度としたルーレット選択によって決定される. 類似度 S_{ij} は, 各エージェントの特徴ベクトル $\mathbf{v}^i, \mathbf{v}^j$ を用いて, 以下の式で定義する.

$$S_{ij} = \frac{\mathbf{v}^i \cdot \mathbf{v}^j}{F} \quad (5)$$

ここで, F はエージェントの特徴ベクトルの長さを表す. これにより, お互いに共通点の多いユーザ程, SNS 内で友達になる可能性が高いことを表現する.

2.5 シミュレーションの手法

シミュレーションの各ターンで, エージェントは以下の行動を順番に行う. すべてのエージェントが行動を終えた時点で2.4節のモデルに従ってネットワークの構築を行う.

ログイン

エージェント a_i はログイン頻度 l_i に従って, ログインする. ログインした場合は, 以下の「コメントの確認」, 「隣接エージェントの日記の確認」, 「コメントの記入」, 「日記の記入」を順番に行う. ログインしなかった場合は, その時点でこのターンの行動を終了する.

コメントの確認

自分が書いた日記に対してのコメントがあるかどうかを確認する. コメントの有無に応じてターン t における日記更新頻度 $d_i(t)$ を以下の式に従って更新する.

$$d_i(t) = \begin{cases} d_i(t-1) - \rho & (\text{新規コメントなし}) \\ d_i(t-1) + \rho n_i / k_i & (\text{新規コメントあり}) \end{cases} \quad (6)$$

ここで, ρ は日記更新頻度変化率を表す定数である. また n_i はエージェント a_i の日記に対する新規コメントの件数を表し, k_i はエージェント a_i の次数を表す.

これによって, 日記に対するコメントが多ければ多いほど, エージェントがそれに満足して日記更新頻度が高くなる様子を表現する.

隣接エージェントの日記の確認

隣接エージェントが日記を更新しているかどうかを確認する. 新着日記の有無に応じてターン t におけるログイン頻度 $l_i(t)$ を以下の式に従って更新する.

$$l_i(t) = \begin{cases} l_i(t-1) - \lambda & (\text{新着日記なし}) \\ l_i(t-1) + \lambda n_i / k_i & (\text{新着日記あり}) \end{cases} \quad (7)$$

ここで, λ はログイン頻度変化率を表す定数である. また n_i は新着日記の件数を表し, k_i はエージェント a_i の次数を表す.

これによって, 周囲のエージェントが日記を多く書いているほど, その日記を読むためにログイン頻度が高くなる様子を表現する.

コメントの記入

一般に, 仲のよい友人の日記ほどコメントを書く可能性が高いと考えられる. そこで, 行動中のエージェント a_i の隣接エージェント $a_j (i \neq j)$ が新しい日記を書いている場合, 以下の確率 $p_c(a_i, a_j)$ に従ってコメントを記入する.

$$p_c(a_i, a_j) = f_{ij} \quad (8)$$

これは, $f_{ij} = 0.6$ であれば60%の確率でコメントを記入することを表す.

日記の記入

日記更新頻度 $d_i(t)$ に従い, 日記を更新する. 日記を更新しない場合は, そのままエージェント a_i の当該ターンにおける行動は終了する.

3. 妥当性検証実験

3.1 実験目的

本論文で提案したモデルの妥当性を確認するために、名古屋大学大学院情報科学研究科において実際に運用されている学内SNS, Academic Community System[2] (以下 ACS と略記) の実データとの比較を行った。検証実験では、提案モデルによって生成したネットワークを ACS のネットワークと比較する。これらのネットワークが類似していれば、提案モデルは SNS のモデルとして妥当であると考えられる。また、比較対象として "Random Graph [5]", "CNN モデル [3]", "BA モデル [6]" によって生成したネットワークと、ACS のネットワークをそれぞれ比較する。

ACS は 2006 年 1 月からサービスを開始した。また、2006 年 12 月 15 日時点で ACS を利用しているユーザ数は 364 人、リンク数は 456 であった。ただし、リンクを 1 本も持たないユーザ (201 人) が存在した。残りのユーザは大小 2 つの連結成分に分割されており、大きな連結成分に含まれるユーザ (158 人)、小さな連結成分に含まれるユーザ (5 人) が存在していた。そこで、以下では ACS のネットワークとして最大連結成分 (ユーザ数 $N=158$, リンク数 $m=452$) を取り出し利用する。

また、サービス開始から 2006 年 12 月 15 日まで 350 日間のユーザ数の推移は図 2 に示す通りであった。

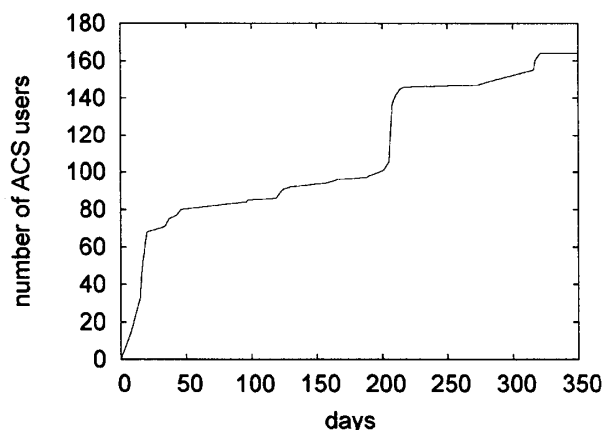


図 2: ユーザ数の推移 (ACS 実データ)

3.2 実験手法

提案モデルによってネットワークを構築し、ACS のネットワークと比較する。また、Random Graph, CNN モデル, BA モデルによって生成したネットワークと、ACS のネットワークの比較をそれぞれ行う。

ネットワーク構造の比較の際、評価指標には、平均経路長、クラスタ係数、次数分布を用いる。

シミュレーションでは、図 2 と同様にユーザ数を推移させる。モデルのパラメータは、SNS 利用者を対象とした調査 [1] を参考に設定した。シミュレーション及びモデルの詳細設定をそれぞれ表 1, 表 2 に示す。シミュレーション期間は、2006 年 1 月 1 日から 2006 年 12 月 15 日まで期間を想定し、350 ターン (350 日) とした。また、

同条件のシミュレーションを 50 回行い、その結果の平均値を評価値とした。

表 1: シミュレーションの詳細設定

シミュレーション期間	350 ターン
シミュレーション回数	50 回

表 2: モデルの詳細設定

ログイン頻度分布の平均 μ_l	0.45
ログイン頻度分布の標準偏差 σ_l	0.10
日記頻度分布の平均 μ_d	0.25
日記頻度分布の標準偏差 σ_d	0.10
特徴ベクトル長 F	256
特徴ベクトル初期化用パラメータ B	8
好感度変化率 α	0.90
好感度下限値 β	0.10
日記更新頻度変化率 ρ	0.04
ログイン頻度変化率 λ	0.02
好感度の決定定数 T	14
SNS 活用度の決定定数 H	14

3.3 実験結果

平均経路長, クラスタ係数

表 3 に、ACS 及び提案モデルによって構築されたネットワークの平均経路長, クラスタ係数を示す。また、Random Graph, BA モデル, CNN モデルによって構築したネットワークのデータを比較として示す。ここで、表中の N , m , L , C はそれぞれエージェント数, リンク数, 平均経路長, クラスタ係数を表す。

表 3: ネットワーク構造の比較

ネットワーク構造	N	m	L	C
ACS(実データ)	158	904	3.33	0.449
提案モデル	158	900	3.42	0.452
Random Graph	158	926	3.48	0.037
BA モデル	158	930	2.80	0.099
CNN モデル	158	900	4.00	0.280

平均経路長, クラスタ係数ともに、提案モデルが ACS の実データと最も近い。これにより、平均経路長, クラスタ係数の観点からは提案モデルがもっとも忠実に ACS のネットワークの構造を表現出来ていると言える。

次数分布

表 4 にそれぞれの次数分布のベキ指数を示す。

表 4 より、提案モデルが ACS の実データに最も近い値を示していることがわかる。また、CNN モデルのベキ指数も ACS の実データと比較的近い値である。BA モデルのベキ指数は、ACS の実データのベキ指数と大きく異なっている。これは、次数分布の低次領域における傾向が異なるためであると考えられる。以上より、提案

表 4: ベキ指数の比較

ネットワーク構造	ベキ指数 γ
ACS 実データ	-1.166
提案モデル	-1.175
BA モデル	-1.628
CNN モデル	-1.106

モデルはベキ指数の観点からも、ACS ネットワークに近い値を示していることがわかる。なお、参考として図 3 に ACS 実ネットワークの度数分布を、図 4 に提案モデルによって構築したネットワークの度数分布の一例を、それぞれ両対数軸でプロットしたものを示す。

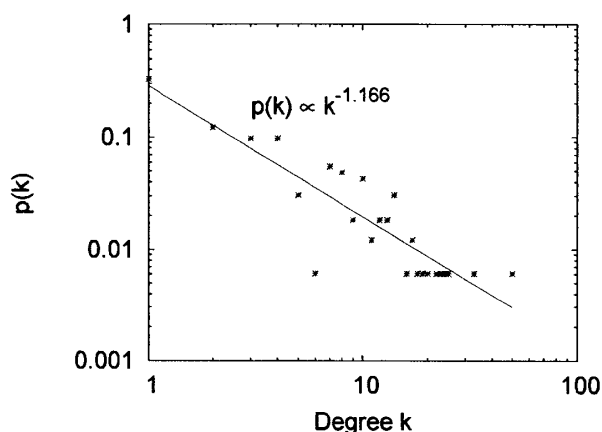


図 3: ACS ネットワークの度数分布 (実データ)

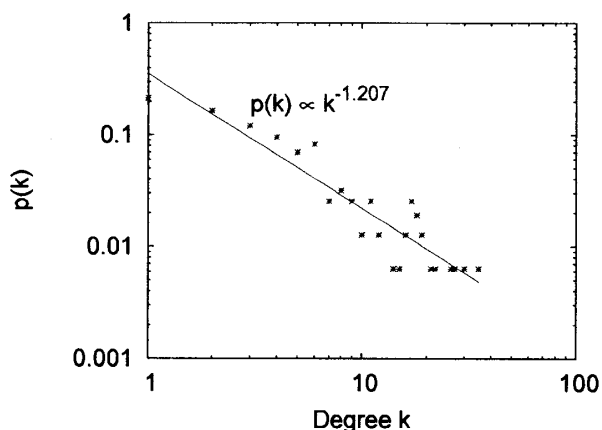


図 4: 提案モデルの度数分布

以上、平均経路長、クラスタ係数、度数分布より、提案モデルが ACS のネットワーク構造をもっとも忠実に表現していることが確認され、モデルの妥当性が確認された。

4. まとめ

本研究では、SNS におけるユーザの振る舞いを表現する SNS コミュニケーションモデルを提案した。提案モデルでは、ユーザが日記及びコメントを通じてコミュニケーションを行う状況表現した。さらに、そのコミュニケーションの活発さや、ユーザ間の類似度の高さに応じてネットワークが形成される過程を表現した。

本モデルによって構築された SNS ネットワークと、名古屋大学学内 SNS として運用されている Academic Community System (ACS) のデータとの比較を行い、提案モデルの妥当性の検証を行った。その結果、本モデルによって ACS のネットワークを表現可能であることを確認した。

今後の課題としては、本論文で提案したモデルを用いて、SNS 上で起こる様々な現象を分析することが挙げられる。具体的には、SNS を活発化させるための手法の分析や、SNS 上での情報伝播の様子などの分析などが考えられる。また、本モデルを ACS 以外の SNS へ適用し、モデルの妥当性及び一般性を検証することも今後の課題の一つである。

謝辞

本研究を行うにあたり、Academic Community System のデータを提供して下さった名古屋大学大学院工学研究科の河口信夫准教授、名古屋大学大学院情報科学研究科の高井一輝氏に感謝します。

参考文献

- [1] ブログ・SNS 利用者の実像～人々は何を求めているのか～. NEC 総研, 2006.
- [2] 高井一輝, 河口信夫. 多様な人間関係を表現可能なアカデミックコミュニティシステム. *The 20th Annual Conference of the Japanese Society for Artificial Intelligence*, 2006.
- [3] A. Vázquez. Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations. *Physical Review E*, Vol. 67, No. 5, p. 56104, 2003.
- [4] G. Bianconi and A.L. Barabasi. Competition and multiscaling in evolving networks. *Europhysics Letters*, Vol. 54, No. 4, pp. 436–442, 2001.
- [5] P. Erdős and A. Rényi. On the evolution of random graphs. *Bulletin of the Institute of International Statistics*, Vol. 38, pp. 343–347, 1961.
- [6] A.L. Barabási and R. Albert. Emergence of Scaling in Random Networks. *Science*, Vol. 286, No. 5439, p. 509, 1999.