

ソーシャルネットワークにおける 潜在特徴を考慮した情報拡散モデル

吉川 友也^{1,a)} 岩田 具治^{2,b)} 澤田 宏^{3,c)}

概要：オンラインコミュニティにおいて、ユーザは流行しているトピックに対する意見や噂といった様々な情報を共有している。このコミュニティにおける将来のトレンド予測を可能にするために、情報拡散モデルとそれのパラメータ推定法を開発することが重要である。この論文では、潜在的なコミュニティ間の影響を捉えることによって、効率的に伝播確率を推定できる潜在特徴独立カスケードモデルを提案する。具体的には、提案モデルは各ノードに情報の送信者としての潜在特徴と受信者としての潜在特徴を持つと仮定する。実験では、提案モデルが従来手法よりも精度高く伝播確率を推定できることを示すとともに、ユーザの影響度推定に対しても有効であることを示す。

1. はじめに

オンラインコミュニティ上では、流行しているトピックに対する意見や噂などの情報がクチコミによって個人間で共有される。オンラインコミュニティでのユーザの行動から、株価 [12] や選挙の得票数 [1] が予測できることから、ソーシャルネットワーク上で共有される情報から将来の流行を予測する研究が近年盛んに行われている [2], [4], [9]。

Kempe ら [6] によって提案された Independent Cascade Model (ICM) は情報拡散現象を説明するモデルとして良く研究されており、影響最大化/最小化問題を解くため [6], [7] や、影響力のあるノードを発見するため [5], [8] に使われている。ICM は、情報がノードからノードへと伝わる過程を表現する確率モデルであり、モデルの挙動は各リンクに付随する伝播確率によって決まる。従って、情報拡散現象をシミュレーションするためには、全てのリンクの伝播確率を正しく推定することが重要である。

ICM の伝播確率を推定するための方法はいくつか提案されている [3], [10], [11]。これらの方法を使うことで伝播確率は推定できるが、観測データが少ない時、汎化性能の低いモデルを学習する可能性がある。例えば、ネットワー

ク全体を通じて情報拡散があまり起こっていない場合や、ネットワークの一部で情報拡散はよく起こるが、他の一部ではあまり起こらない場合が考えられる。そのような場合は悪いパラメータ推定結果に繋がる可能性がある。

本論文では、効率的に伝播確率を推定できるように設計した Latent Feature Independent Cascade Model (LFICM) を提案する。LFICM では、各ノードは潜在的に情報の送信者としての特徴と、受信者としての特徴を持つと仮定する。また、伝播確率は送受信両ユーザの潜在特徴に依存して決まるとする。潜在特徴をモデルに取り入れることにより、提案モデルは通常の ICM と比較して少ないパラメータ数になり、高い汎化性能でパラメータ推定が可能になる。LFICM のパラメータ推定は EM アルゴリズムに基づいて行う。実験では、実ネットワークデータと人工で作成した拡散データを使い、提案モデルが既存手法よりも精度高く伝播確率を推定できることを示す。また、提案モデルで推定した伝播確率を使うことによって、真の影響度と同じ振る舞いをする影響度を推定できることを示す。

2. 提案法

2.1 モデル

情報の集合 I が有向ネットワーク $G = (\mathcal{V}, \mathcal{E})$ 上で拡散するとする。ここで、 \mathcal{V} はソーシャルネットワークにおけるユーザに対応するノード、 $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ はノード間の関係を表すリンクである。 $B(v)$ をノード v の親ノード集合、 $F(v)$ をノード v の子ノード集合と定義する。

観測データとして、情報 $i \in I$ を発信したノードとその時刻の組からなる拡散系列 d_i が得られる。

¹ 奈良先端科学技術大学院大学
Nara Institute of Science and Technology, Takayama, Ikoma,
Nara 630-0192, Japan
² 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories
³ 日本電信電話株式会社 NTT サービスエボリューション研究所
NTT Service Evolution Laboratories
a) yuya-y@is.naist.jp
b) iwata.tomoharu@lab.ntt.co.jp
c) sawada.hiroshi@lab.ntt.co.jp

$$d_i = \{(v_{i1}, t_{v_{i1}}), (v_{i2}, t_{v_{i2}}), \dots, (v_{iL_i}, t_{v_{iL_i}})\}.$$

なお、情報を発信したノードの状態をアクティブ、発信していないノードの状態を非アクティブと表現する。ここで、アクティブノードは親ノードのいずれか一つからのみ影響を受け、我々はその影響を与えた親ノードを観測することはできないとする。簡便のため、 $D = \{d_1, d_2, \dots, d_M\}$ と表記する。

LFICM は Independent Cascade Model (ICM) [6] に基づいて情報が広がることを仮定している。ICM は各リンク $(u, v) \in \mathcal{E}$ に対して、伝播確率 $\kappa_{uv}, 0 \leq \kappa_{uv} \leq 1$ と時間遅れパラメータ $r_{uv} > 0$ が割り当てられている。簡単のため、本論文では $r_{uv} = r$ と設定する。

ICM の情報拡散過程は Algorithm 1 で示す。

Algorithm 1 ICMGENERATOR

Require: ネットワーク \mathcal{G} , 情報源ノード集合 S , 時間遅れパラメータ r , 伝播確率 κ

```

1:  $d_m \leftarrow \{(u, 0) | u \in S\}$ 
2: repeat
3:    $(u, t_{mu}) \leftarrow \min_{t_{mx}} \{(x, t_{mx}) | (x, t_{mx}) \in d_m, x \in S\}$ 
4:   for  $v \in \{u \text{ の子ノードで、非アクティブなノード}\}$  do
5:     if  $u$  が  $\kappa_{uv}$  の確率で  $v$  へ拡散成功 then
6:        $t_{mv} \leftarrow t_{mu} + \delta$ , where  $\delta \sim \text{Exponential}(\delta | r_{uv})$ 
7:        $d_m \leftarrow d_m \cup \{(v, t_{mv})\}$ 
8:        $S \leftarrow S \cup \{v\}$ 
9:     end if
10:  end for
11:   $S \leftarrow S \setminus \{u\}$ 
12: until  $S = \emptyset$ 
13: return  $d_m$ 

```

LFICM では、各ノード $u \in \mathcal{V}$ において、2 種類の潜在特徴ベクトル $\mathbf{x}_u \in \mathbb{R}^K, \mathbf{y}_u \in \mathbb{R}^K$ を仮定する。そして、ノード u から v への伝播確率 κ_{uv} は以下のように計算する。

$$\kappa_{uv} = f(\mathbf{x}_u, \mathbf{y}_v, \gamma) = (1 + \exp(-\mathbf{x}_u^\top \mathbf{y}_v - \gamma))^{-1}, \quad (1)$$

ここで、 γ はノードに依存しないバイアスパラメータである。関数 f はシグモイド関数であり、 $0 \leq \kappa_{uv} \leq 1$ が保証される。 \mathbf{x}_u と \mathbf{y}_u はそれぞれ、情報の送信者としての特徴と受信者としての特徴を表現しており、 $\mathbf{x}_u^\top \mathbf{y}_v$ が大きい値ならば、伝播確率 κ_{uv} もまた大きい確率を持つ。

ノード $u \in \mathcal{V}$ の潜在特徴ベクトル $\mathbf{x}_u, \mathbf{y}_u$ はそれぞれ以下の K 次元正規分布に従う。

$$\mathbf{x}_u \sim \mathcal{N}(\mathbf{x}_u | \mathbf{0}, \sigma_X^2 \mathbf{1}), \quad \mathbf{y}_u \sim \mathcal{N}(\mathbf{y}_u | \mathbf{0}, \sigma_Y^2 \mathbf{1}), \quad (2)$$

ここで、 $\mathbf{0}$ と $\mathbf{1}$ はそれぞれ、 K 次元の零ベクトルと $K \times K$ の単位行列を表す。情報源ノード $S_i \subseteq \mathcal{V}$ と上記で計算された伝播確率の下、拡散系列 d_i は Algorithm 1 で示す情報拡散過程に基づいて生成される。

$$d_i \sim \text{ICM}(\{\kappa_{uv}\}_{(u,v) \in \mathcal{E}}, r, \mathcal{G}, S_i). \quad (3)$$

2.2 パラメータ推定

$\Delta_{uv}^{(i)} = t_{iv} - t_{iu}$ と定義する。また、情報 i におけるアクティブノード集合 $C_i = \{u | (u, t) \in d_i\}$ 、時刻 t までのアクティブ集合 $C_i(t) = \{u | (u, t') \in d_i, t' < t\}$ とする。式 (2), (3) の仮定から、潜在特徴 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_{|\mathcal{V}|}], \mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_{|\mathcal{V}|}]$ の事後確率は以下で与えられる。

$$\begin{aligned}
P(\mathbf{X}, \mathbf{Y} | D, \gamma, r, \mathcal{G}) & \quad (4) \\
&= \prod_{i \in \mathcal{I}} \prod_{v \in C_i} \prod_{u \in B(v) \cap C_i(t_{iv})} p_{u \rightarrow v}^{(i)} \sum_{u \in B(v) \cap C_i(t_{iv})} \frac{p_{u \rightarrow v}^{(i)}}{p_{u \rightarrow v}^{(i)}} \\
&\quad \times \prod_{w \in F(v) \setminus C_i} (1 - \kappa_{vw}) \times P(\mathbf{X} | \sigma_X) P(\mathbf{Y} | \sigma_Y)
\end{aligned}$$

ここで、 $p_{u \rightarrow v}^{(i)}$ は情報 i においてノード u からノード v に情報伝播が成功する確率、 $p_{u \rightarrow v}^{(i)}$ はノード u からノード v への情報伝播が失敗する確率を表し、以下のように書ける。

$$p_{u \rightarrow v}^{(i)} = \kappa_{uv} r \exp(-r \Delta_{uv}^{(i)}), \quad (5)$$

$$p_{u \rightarrow v}^{(i)} = \kappa_{uv} \exp(-r \Delta_{uv}^{(i)}) + 1 - \kappa_{uv}. \quad (6)$$

事後確率式 (4) を最大にするパラメータ $\hat{\mathbf{X}}, \hat{\mathbf{Y}}, \hat{\gamma}, \hat{r}$ は、EM アルゴリズムに基づく反復アルゴリズムによって推定する。なお、表記上の混乱を避けるため、 \bar{x} と書くときは、パラメータ x の現在の推定値を表すこととする。

EM アルゴリズムでは、式 (4) の代わりに、式 (4) の下限となる以下の関数 Q を最大化する。

$$\begin{aligned}
Q(\mathbf{X}, \mathbf{Y}, \gamma, r; \bar{\mathbf{X}}, \bar{\mathbf{Y}}, \bar{\gamma}, \bar{r}) & \quad (7) \\
&= \sum_{i \in \mathcal{I}} \sum_{v \in C_i} \left[\sum_{u \in B(v) \cap C_i(t_{iv})} (\bar{\xi}_{uv}^{(i)} \log \kappa_{uv} \right. \\
&\quad \left. + (1 - \bar{\xi}_{uv}^{(i)}) \log(1 - \kappa_{uv}) + \bar{q}_{uv}^{(i)} \log r + \bar{\xi}_{uv}^{(i)} r \Delta_{uv}^{(i)} \right) \\
&\quad \left. + \sum_{w \in F(v) \setminus C_i} \log(1 - \kappa_{vw}) \right] \\
&\quad + \log P(\mathbf{X} | \sigma_X) + \log P(\mathbf{Y} | \sigma_Y).
\end{aligned}$$

ここで、

$$q_{uv}^{(i)} = \frac{p_{u \rightarrow v}^{(i)} / p_{u \rightarrow v}^{(i)}}{\sum_{u' \in B(v) \cap C_i(t_{iv})} p_{u' \rightarrow v}^{(i)} / p_{u' \rightarrow v}^{(i)}}, \quad (8)$$

$$\eta_{uv}^{(i)} = \frac{\kappa_{uv} \exp(-r \Delta_{uv}^{(i)})}{\kappa_{uv} \exp(-r \Delta_{uv}^{(i)}) + (1 - \kappa_{uv})}, \quad (9)$$

$$\xi_{uv}^{(i)} = q_{uv}^{(i)} + (1 - q_{uv}^{(i)}) \eta_{uv}^{(i)}. \quad (10)$$

そして、関数 Q を最大にする $\mathbf{X}, \mathbf{Y}, \gamma$ は、以下の一階微分の値を使った L-BFGS 法によって推定される。

$$\begin{aligned}
\frac{\partial Q(\mathbf{X}, \mathbf{Y}, \gamma, r; \bar{\mathbf{X}}, \bar{\mathbf{Y}}, \bar{\gamma}, \bar{r})}{\partial \mathbf{x}_u} & \quad (11) \\
&= \sum_{i \in \mathcal{I}} \left(\sum_{v \in F(u) \cap C_i} (\bar{\xi}_{uv}^{(i)} - f(\mathbf{x}_u, \bar{\mathbf{y}}_v, \bar{\gamma})) \bar{\mathbf{y}}_v \right. \\
&\quad \left. - \sum_{w \in F(u) \setminus C_i} f(\mathbf{x}_u, \bar{\mathbf{y}}_w, \bar{\gamma}) \bar{\mathbf{y}}_w \right) - \frac{1}{\sigma_X^2} \mathbf{x}_u
\end{aligned}$$

$$\begin{aligned} & \frac{\partial Q(\mathbf{X}, \mathbf{Y}, \gamma, r; \bar{\mathbf{X}}, \bar{\mathbf{Y}}, \bar{\gamma}, \bar{r})}{\partial \mathbf{y}_v} \quad (12) \\ &= \sum_{i \in \mathbf{I}} \left(\sum_{u \in B(v) \cap C_i(t_{iv})} (\bar{\xi}_{uv}^{(i)} - f(\bar{\mathbf{x}}_u, \mathbf{y}_v, \bar{\gamma})) \bar{\mathbf{x}}_u \right. \\ & \quad \left. - \sum_{s \in B(v) \cap C_i, v \notin C_i} f(\bar{\mathbf{x}}_s, \mathbf{y}_v, \bar{\gamma}) \bar{\mathbf{x}}_s \right) - \frac{1}{\sigma_Y^2} \mathbf{y}_v \end{aligned}$$

$$\begin{aligned} & \frac{\partial Q(\mathbf{X}, \mathbf{Y}, \gamma, r; \bar{\mathbf{X}}, \bar{\mathbf{Y}}, \bar{\gamma}, \bar{r})}{\partial \gamma} \quad (13) \\ &= \sum_{i \in \mathbf{I}} \left(\sum_{v \in C_i} \sum_{u \in B(v) \cap C_i(t_{iv})} (\bar{\xi}_{uv}^{(i)} - f(\bar{\mathbf{x}}_u, \bar{\mathbf{y}}_v, \gamma)) \right. \\ & \quad \left. - \sum_{v \in C_i} \sum_{s \in F(v) \setminus C_i} f(\bar{\mathbf{x}}_s, \bar{\mathbf{y}}_v, \gamma) \right) \end{aligned}$$

また、時間遅れパラメータ r は以下の閉形式で更新する。

$$r = \frac{\sum_{i \in \mathbf{I}} \sum_{v \in C_i} \sum_{u \in B(v) \cap C_i(t_{iv})} \bar{q}_{uv}^{(i)}}{\sum_{i \in \mathbf{I}} \sum_{v \in C_i} \sum_{u \in B(v) \cap C_i(t_{iv})} \bar{\xi}_{uv}^{(i)} \Delta_{uv}^{(i)}}. \quad (14)$$

パラメータ推定の手続きをまとめると、Algorithm 2 となる。

Algorithm 2 LFICMESTIMATOR

Require: ネットワーク \mathcal{G} , 拡散系列 \mathbf{D} , 特徴ベクトルの長さ K , ハイパーパラメータ σ_X, σ_Y

- 1: $\mathbf{X}, \mathbf{Y}, \gamma, r$ を初期化
- 2: **repeat**
- 3: **E** ステップ: 式 (8),(9),(10) を使って, $q_{uv}^{(i)}, \eta_{uv}^{(i)}, \xi_{uv}^{(i)}$ を更新
- 4: **M** ステップ: 式 (11),(12),(13) を使った準ニュートン法によって $\mathbf{X}, \mathbf{Y}, \gamma$ を更新。式 (14) を使って r を更新。
- 5: **until** 関数 (7) の値の上昇が収束
- 6: **return** $\mathbf{X}, \mathbf{Y}, \gamma, r$

3. 実験

3.1 実験設定

ネットワーク構造: 3種類の実ネットワーク構造 BLOG, ENRON, MIXI を使用した。ページ数の制限のため、詳細は割愛する。表 1 は各データセットの統計量を示す。

拡散系列の生成: 今回の実験では、拡散系列 d_i は提案モデルによって人工的に生成した。まず、各データセットの σ_X, σ_Y の設定の下で潜在特徴ベクトル \mathbf{X}, \mathbf{Y} を生成し、伝播確率を計算する。その後、Algorithm 1 によって d_m を生成する。なお、提案法のパラメータ推定の頑健性を確認するために、各データセットのリンク数 $|\mathcal{E}|$ に対して、拡散系列の数 $|\mathbf{I}|$ と平均の拡散系列長 $\text{avg.}|d_i|$ が小さくなるように拡散系列を生成した。表 2 は拡散系列生成時の設定の一覧を示す。情報源ノード集合 \mathcal{S} の Random は、ノード集合中からランダムに 1 つ情報源ノードとしたことを意味する。

対抗手法: 2種類のパラメータ推定法 SaitoICM [10],

表 1 ネットワークデータの統計量

	$ \mathcal{V} $	$ \mathcal{E} $	$ \mathbf{I} $	avg. $ d_i $
BLOG	12,047	79,920	200	207.1
ENRON	36,692	367,662	200	203.4
MIXI	80,608	571,136	1,500	8.6

表 2 拡散系列生成のパラメータ設定

	K	γ	r	\mathcal{S}	σ_X	σ_Y
BLOG	5	-3.5	10.0	Random	1.0	1.0
ENRON	5	-6.5	1.0	Random	1.0	1.0
MIXI	7	-6.5	1.0	Random	1.0	1.0

SimpleICM を対抗手法とする。SimpleICM はネットワーク全体を一つの伝播確率で表現するモデルである。二つ目は、全てのリンクの伝播確率を同一の値として推定する方法である。

3.2 パラメータ推定の精度

図 1 は、各データセットにおける提案法とベースラインの伝播確率の推定精度の結果を示す。縦軸は真の伝播確率と推定した伝播確率との RMSE、横軸は推定時の特徴ベクトルの長さ K である。RMSE は小さいほど良い推定精度を表す。なお、各推定法による推定結果は初期値に依存するため、同じ情報拡散のログに対して 5 回パラメータ推定を行って評価した。BLOG と ENRON は、共に $K = 1$ の時、SimpleICM に負けている。これは、真の K^* が 5 で拡散系列を生成しているために、 $K = 1$ では観測データの特徴を上手く捉えきれなかったのが原因だと考えられる。それ以外の K では、LFICM が最も良い精度である。MIXI は、LFICM の推定結果の標準偏差が比較的大きいのが特徴的である。これは、MIXI は平均の $|d_m|$ が他と比べてかなり小さい設定でデータを生成しているため、パラメータ推定に十分なデータ数が得られていないことが原因と考えられる。しかし、 $K = 3, 5, 7$ では、LFICM が最も良い精度であった。

3.3 影響度推定精度

ノードから発信される情報によってアクティブになるノードの総数の平均をそのノードが持つ影響度と定義する。情報拡散モデルを使った影響度推定の方法として、[13] が既に提案されている。この方法による推定性能は、推定した情報拡散モデルのパラメータの正確性に大きく依存する。我々は、[13] の方法を Algorithm 3 のように単純化し、提案法 LFICM で推定したパラメータを用いた場合の影響度の予測性能を評価する。今回の実験では、試行回数 T は 50 に設定した。

表 3 は、真の影響度と各手法で推定した影響度のピアソン相関係数の一覧である。この問題設定では、ピアソン相関係数は 1 に近いほど良い性能と見ることが出来る。こ

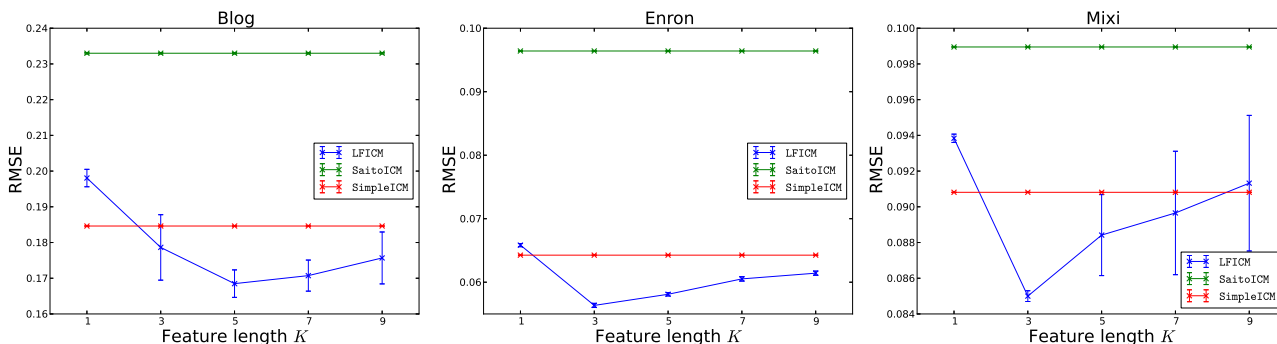


図 1 真の伝播確率と推定した伝播確率の RMSE

表 3 真の影響度と推定した影響度のピアソン相関係数 . 太字は各データセットにおいて最も良い結果を示す

	LFICM ($K = 1$)	LFICM ($K = 3$)	LFICM ($K = 5$)	LFICM ($K = 7$)	LFICM ($K = 9$)	SaitoICM	SimpleICM
BLOG	0.360	0.734	0.848	0.828	0.769	0.753	0.827
ENRON	0.442	0.730	0.703	0.669	0.657	0.539	0.656
MIXI	0.174	0.542	0.545	0.556	0.484	0.468	0.382

Algorithm 3 INFLUENCEPREDICTOR

Require: ネットワーク \mathcal{G} , 情報源ノード集合 S , 時間遅れパラメータ r , 伝播確率 $\{\kappa_{uv}\}_{(u,v) \in \mathcal{E}}$, 試行回数 T

```

1:  $influence \leftarrow 0$ 
2: for  $m \leftarrow 1, 2, \dots, T$  do
3:    $\mathbf{d}_m \leftarrow \text{ICMGENERATOR}(\mathcal{G}, S, r, \{\kappa_{uv}\}_{(u,v) \in \mathcal{E}})$  // Alg. 1
4:    $influence \leftarrow influence + \frac{1}{T} |\mathbf{d}_m|$ 
5: end for
6: return  $influence$ 

```

の表を見ると、いずれのデータセットにおいても、提案法 LFICM で推定したパラメータは、ベースラインで推定したパラメータよりも良い影響度推定の結果を導くことが分かる。

4. おわりに

本稿では、独立カスケードモデルのパラメータをより正確に推定するために、ノードの潜在特徴を考慮した情報拡散モデルを提案した。

参考文献

[1] Franch, F.: 2010 UK Election Prediction with Social Media, *Journal of Information Technology & Politics*, Vol. 10, No. 1, pp. 57–71 (2013).
[2] Galuba, W. and Aberer, K.: Outtweeting the Twitterers-Predicting Information Cascades in Microblogs, *Proceedings of the 3rd Conference on Online Social Networks* (2010).
[3] Gruhl, D., Guha, R., Liben-Nowell, D. and Tomkins, A.: Information Diffusion through Blogspace, *Proceedings of the 13th International Conference on World Wide Web*, pp. 491–501 (2004).
[4] Iribarren, J. and Moro, E.: Impact of Human Activity Patterns on the Dynamics of Information Diffusion, *Physical Review Letters*, Vol. 103, No. 3, pp. 8–11

(2009).
[5] Kempe, D., Kleinberg, J. and Tardos, E.: Influential Nodes in a Diffusion Model for Social Networks, *Automata, Languages and Programming*, Vol. 3580, pp. 1127–1138 (2005).
[6] Kempe, D., Kleinberg, J. and Tardos, E.: Maximizing the Spread of Influence through a Social Network, *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 137 (2003).
[7] Kimura, M., Saito, K. and Motoda, H.: Blocking Links to Minimize Contamination Spread in a Social Network, *ACM Transactions on Knowledge Discovery from Data*, Vol. 3, No. 2, pp. 1–23 (2009).
[8] Kimura, M., Saito, K. and Nakano, R.: Extracting Influential Nodes for Information Diffusion on a Social Network, *Proceedings of the National Conference on Artificial Intelligence*, Vol. 22, No. 2, p. 1371 (2007).
[9] Kuo, T.-T., Hung, S.-C., Lin, W.-S., Lin, S.-D., Peng, T.-C. and Shih, C.-C.: Assessing the Quality of Diffusion Models Using Real-World Social Network Data, *2011 International Conference on Technologies and Applications of Artificial Intelligence*, pp. 200–205 (2011).
[10] Saito, K., Kimura, M., Ohara, K. and Motoda, H.: Learning Continuous-Time Information Diffusion Model for Social Behavioral Data Analysis, *First Asian Conference on Machine Learning*, pp. 322–337 (2009).
[11] Saito, K., Ohara, K., Yamagishi, Y. and Kimura, M.: Learning Diffusion Probability Based on Node Attributes in Social Networks, *19th International Symposium, ISMIS 2011*, pp. 153–162 (2011).
[12] Zhang, X., Fuehres, H. and Gloor, P. A.: Predicting Stock Market Indicators Through Twitter “I hope It Is Not as Bad as I Fear”, *The 2nd Collaborative Innovation Networks Conference*, Vol. 26, pp. 55–62 (2011).
[13] 吉川友也, 齊藤和巳, 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, *電子情報通信学会論文誌. D, 情報・システム*, Vol. 94, No. 11, pp. 1899–1908 (2011).