

ネット上のコミュニティにおける情報の 評判計算アルゴリズム EigenRumor

谷本 直人¹

藤村 考²

ネット・コミュニティにおいて提供される様々な情報の評判を定量化する方法を述べる。本稿で提案する EigenRumor アルゴリズムは、コミュニティの参加者による情報の提供と、提供された情報に対する参加者による評価という2種類の行為を参加者から情報へのリンクとみなし、リンク分析によりそのネット・コミュニティでの情報の評判をスコアリングする。本アルゴリズムは、コミュニティ参加者の貢献度も同時に算出するため、コミュニティの中の「良い」参加者を特定するためにも有効である。また、評判の算出で重要な「さくら攻撃」に対する耐性について評価した結果を報告する。

The EigenRumor Algorithm for Calculating Reputation of Information Resources in Electronic Communities

Naoto Tanimoto¹

Ko Fujimura²

This paper describes a method for quantifying the degree of reputation for information resources provided in Electronic communities. The EigenRumor algorithm, proposed here, calculates the reputation scores based on a link analysis approach by considering information provisioning and information evaluation actions as links from participants to information resources. This algorithm also calculates the contribution level of participants in the electronic community and it is used to identify "good" participant as well as "good" information. The robustness of the algorithms against "ballot stuffing" attack is also discussed.

1 はじめに

近年、ネットワーク技術の進歩、ネットワーク利用料金の低価格化により、ネットワーク利用者数が増加している。それに伴い、ネットワーク上で利用者が情報交換や情報提供するネット・コミュニティが形成され、活発化しつつある。

ネット・コミュニティでは玉石混交な多数の情報が提供されるため、コミュニティの参加者同士で情報を評価し、その評価を集約して評判として参加者に提示し、参加者への情報選択の指標を与える評判システムが提案されているが、評判システムが有効に機能するためには、質の良い評価を集めて信頼性の高い評判を計算する必要がある。

本稿では、コミュニティ上での参加者による情報の提供と、情報の評価という2つの行為を、参加者から情報へのリンクとして表し、そのリンクを解析することにより情報の評判スコア、参加者の情報提供スコアと情報評価スコアを計算する EigenRumor アルゴリズムを提案する。

EigenRumor アルゴリズムは PageRank[1] や HITS[2]と同様にリンクから生成した行列の固有ベクトルを計算することにより各スコアを計算することを基本とする。ただし、PageRank や HITS はオブジェクト間のリンクからオブジェクトの支持の度合いを示す隣接行列を生成するのに対し、EigenRumor はエージェント(参加者)とオブジェクト(情報)間のリンクから行列を生成することを特徴とする。情報提供スコアは参加者が良い情報を提供する能力を表し、情報評価スコアは参加者の情報評価能力の高さを表す。

ネット上では参加者の特定が困難であるため、結託した参加者は、情報を不正評価することにより評判システムを混乱させ、不正に利益を得ようとするのが想定できる。本稿では、結託した参加者による代表的な攻撃であるさくら評価からの攻撃耐性を検証した。

以後、2 節では前提とする参加者間で情報の提供と評価を行うコミュニティを定義する。3 節ではコミュニティにおける行為をモデル化し、EigenRumor を設計するもととなった仮説と計算アルゴリズムを述べる。4 節では、提案アルゴリズムの攻撃耐性を検証するためのコミュニティを設定し、検証方法を示す。5 節では、さくら評価に対する攻撃耐性の検証結果を示す。6 節では他の既存技術との比較と、提案アルゴリズム

¹ 日本電信電話株式会社 NTT 情報流通プラットフォーム研究所
NIT Information Sharing Platform Laboratories, NIT Corporation

² 日本電信電話株式会社 NTT サイバーソリューション研究所
NIT Cyber Solutions Laboratories, NIT Corporation

他の応用方法、今後の検討課題について述べる。

2 コミュニティ

ネット・コミュニティは、それぞれ何らかの目的を持って形成される。本稿では不特定多数のコミュニティ参加者により情報を共有するコミュニティを前提とする。コミュニティには、その目的により、あるトピックに関する問題について参加者で相互にサポートを行う OKWeb[3]のような「問題解決型コミュニティ」、商品やサービスに対する参加者のレビューを共有するための @cosme[4] や Epinions.com[5] のような「商品・サービス評価コミュニティ」、ユーザ主導による新製品・サービス創造を目的とした、たのみこむ[6]のような「商品・サービス創造コミュニティ」、あるトピックについてユーザ間でのディスカッションや情報共有を行う Slashdot[7] のような「トピックベースコミュニティ」等がある。

これらのコミュニティは目的が異なるため、コミュニティを活性化するために必要な要件も異なる。例えば、問題解決型コミュニティでは、質問者の質問意図に対してよりの確かな回答を得ることや、当コミュニティの方向性にあった質問を集めることであろう。商品・サービス評価コミュニティにおいては、当コミュニティ参加者の大半の賛同を得られるような的確な評価を集めることであろう。

このように各コミュニティでは様々な目的があるが、参加者が行っている行為は「情報提供」と「情報評価」という2つの行為に抽象化できる。例えば、問題解決型コミュニティにおいては、質問や回答を投稿する「情報提供」行為と、質問者が的確な回答内容にポイントを与えたり、質問への参照回数や回答数が「情報評価」行為である考えられる。商品・サービス評価コミュニティにおいては、商品・サービスの評価記事を投稿する「情報提供」行為と、商品・サービスの有用性を評価したり、評価記事に賛同・不賛同を表明する「情報評価」行為が行われていると考えることができる。

上記の分析により、筆者らは、このように情報共有コミュニティから抽出される「情報提供」と「情報評価」の行為を参加者から情報へのリンクとしてとらえ、これらのリンク分析から、コミュニティへの各参加者の貢献をモデル化、スコア化し、さらに参加者の貢献の度合いにより情報提供、情報評価を重み付けすることにより情報の評判をスコア化できないかと考えた。

ただし、リンク分析のみから貢献度を算出するため、コミュニティの活性化に貢献している参加者とは、所属するコミュニティの方向性に合致した「情報提供」あるいは「情報評価」を行っている人と仮定する。ま

た、各参加者にスコアリングするため、コミュニティにおいて参加者を識別できる ID が必要である。次節では、これらの前提および仮定の下で、貢献度をスコアリングし、評判を計算するモデルとアルゴリズムを述べる。

3 EigenRumor アルゴリズム

3.1 コミュニティモデル

参加者(以降、エージェント) m 人、情報(以降、オブジェクト) n 個からなるコミュニティを考える。エージェント i がオブジェクト j を提供したとき、 i から j への「提供リンク」を生成する。コミュニティにおける全提供リンクの状態を、提供行列 $P = [p_{ij}] (i = 1 \dots m, j = 1 \dots n)$ で表す。すなわち、 $p_{ij} = 1$ はエージェント i がオブジェクト j を提供したことを、 0 ならば提供していないことを示す。エージェント i が既に提供されているオブジェクト j の有用性を e_{ij} と評価したとき、 i から j への「評価リンク」を生成する(図 1)。コミュニティにおける全評価リンクの状態を、評価行列 $E = [e_{ij}] (i = 1 \dots m, j = 1 \dots n)$ で表す。 e_{ij} は j に対する i の賛同のレベルを表すが、ここでは $[0, 1]$ の値を取ることにし、 1 が最大の賛同レベルを表すこととする。なお、計算の単純化のために負の値は考慮しない。

なお、この評価値 e_{ij} は良い/悪いといった明示的な評価だけではなく、例えば、質問への回答参照数 k を回答の良さとみなすといった暗示的な評価も含み、この場合は例えば $e_{ij} = 1 - 2^{-k}$ とすることができる。

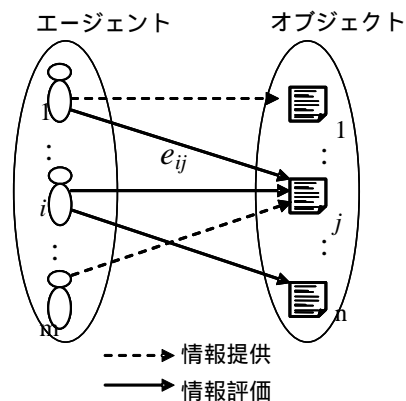


図 1 情報共有コミュニティモデル

3.2 算出するスコア

参加者のコミュニティへの貢献度は、コミュニティの目的により異なるであろうが、本稿では参加者が行った情報提供、情報評価がそのコミュニティ参加者の大半から賛同される方向性と合致している度合いを表

す指標とする。また、情報の評判は、情報がコミュニティの方向性と合致している度合いを示す指標とする。

EigenRumor アルゴリズムでは、エージェントの貢献度を情報提供スコアと情報評価スコアの2つのスコアで表し、各エージェントの行う情報提供と情報評価を、それぞれ情報提供スコアと情報評価スコアで重み付けて情報の評判を計算する。

情報提供スコア (\vec{a})

エージェント i の情報提供の貢献度を表す。コミュニティの方向性と合致する情報提供を行うとスコアが大きくなる。エージェント i の情報提供スコア a_i ($i=1..m$) を要素とするベクトルを \vec{a} と定義する。

情報評価スコア (\vec{h})

エージェント i の情報評価の貢献度を表す。コミュニティの方向性と合致する情報評価を行うとスコアが大きくなる。エージェント i の情報評価スコア h_i ($i=1..m$) を要素とするベクトルを \vec{h} と定義する。

評判スコア (\vec{r})

オブジェクト j の評判を表す。コミュニティの方向性と合致するオブジェクトはスコアが大きくなる。情報提供したエージェントの情報提供スコアと、情報評価したエージェントの情報評価スコアにより重み付けられた評価値の合計で計算する。オブジェクト j の評判スコア r_j ($j=1..n$) を要素とするベクトルを \vec{r} と定義する。

3.3 スコアの相互作用

3.2 で示した3つのスコアの相互作用、EigenRumor アルゴリズムによるスコアの計算方法とその背景となる仮説を説明する。

評判スコア 情報提供スコア、情報評価スコア

次の仮説により、オブジェクトの評判スコアを情報提供を行ったエージェントの情報提供スコアと、投入された評価値を情報評価を行ったエージェントの情報評価スコアにより重み付けした合計で計算する。

仮説 1 良い情報提供エージェントにより提供されたオブジェクトは、コミュニティの方向性に合致する。

仮説 2 良い情報評価エージェントにより支持されたオブジェクトは、コミュニティの方向性に合致する。

情報提供スコア a_i のエージェントにより提供されたオブジェクト j の評判スコアの初期値は $r_j = a_i$ とできる。次に、エージェント i により評価値 e_{ij} の評価が行われると、 i の情報評価スコア h_i により重み付けられた $e_{ij}h_i$ を r_j に加える。つまり、オブジェクト j の評判スコア r_j は次式で計算する：

$$r_j = \alpha a_{prov(j)} + (1 - \alpha) \sum_{i \in eval(j)} e_{ij} h_i \quad (1)$$

ここで、 $prov(j)$ は j を提供したエージェントを、 $eval(j)$ は j を評価したエージェントの集合を表す。は[0,1]の範囲の定数で、情報提供スコアと情報評価スコアの反映の割合をコミュニティの目的により調整する。は時間、あるいは評価数が多くなるにしたがって小さくすることも考えられる。

式(1)を次のように行列表現することもできる：

$$\vec{r} = \alpha P^T \vec{a} + (1 - \alpha) E^T \vec{h} \quad (2)$$

情報提供スコア 評判スコア

次の仮説により、エージェントの情報提供スコアを、エージェントが提供したオブジェクトの評判スコアから計算する。

仮説 3 コミュニティの方向性に合致したオブジェクトを提供するエージェントは、コミュニティへの情報提供の貢献度が大きい。

提供したオブジェクトが、コミュニティとの方向性が合致することにより情報評価を受けると、オブジェクトの評判スコアの増分だけ情報提供スコアを増加する。つまり、エージェント i の情報提供スコア a_i は次式で計算する：

$$a_i = \sum_{j \in prov(i)} r_j \quad (3)$$

ここで、 $prov(i)$ はエージェント i が提供したオブジェクトの集合を表す。

式(3)を次のように行列表現することもできる：

$$\vec{a} = P \vec{r} \quad (4)$$

情報評価スコア 評判スコア

次の仮説により、エージェントの情報評価スコアを、エージェントが評価したオブジェクトの評判スコアから計算する。

仮説 4 コミュニティの方向性に合致したオブジェクトを支持するエージェントは、コミュニティへの情報評価の貢献度が大きい。

評価(支持)したオブジェクトが、他のエージェントにより評価されると、コミュニティの方向性に合致した評価を行ったと考えられるため、その評価値に応じてオブジェクトの評判スコアの増加分だけ情報評価ス

コアを増加する。つまり、エージェント i の情報評価スコア h_i は次式で計算する。

$$h_i = \sum_{j \in \text{eval}(i)} e_{ij} r_j \quad (5)$$

ここで、 $\text{eval}(i)$ はエージェント i が評価したオブジェクトの集合を表す。

式(5)を次のように行列表現することもできる。

$$\vec{h} = E \vec{r} \quad (6)$$

ここで、式(2),(4),(6)は3つのスコアが再帰的な関係となっている。すなわち、式(2)は式(4),(6)から次のように計算する:

$$\begin{aligned} \vec{r} &= \alpha P^T P \vec{r} + (1 - \alpha) E^T E \vec{r} \quad (7) \\ &= (\alpha P^T P + (1 - \alpha) E^T E) \vec{r}, \\ &= S \vec{r} \end{aligned}$$

ここで S は

$$S = (\alpha P^T P + (1 - \alpha) E^T E).$$

S は非負行列であり、HITS[2]と同様に(7)式を繰り返し演算し、各ループで \vec{r} を正規化することにより、 \vec{r} は S の優固有ベクトルに収束する。図2に示すアルゴリズムにより、 \vec{r} と同時に \vec{a} 、 \vec{h} も同時に計算することができる。

```

 $\vec{a}^{(0)} = (1, \dots, 1)^T$ 
 $\vec{h}^{(0)} = (1, \dots, 1)^T$ 
while  $\vec{r}$  change significantly do
   $\vec{r}^{(k)} = \alpha P^T \vec{a}^{(k)} + (1 - \alpha) E^T \vec{h}^{(k)}$ 
   $\vec{r}^{(k+1)} = \vec{r}^{(k)} / \|\vec{r}^{(k)}\|_2$ 
   $\vec{a}^{(k+1)} = P \vec{r}^{(k+1)}$ 
   $\vec{h}^{(k+1)} = E \vec{r}^{(k+1)}$ 
endwhile

```

図2 EigenRumor アルゴリズム

4 アルゴリズムの検証

前節で示したように、EigenRumor アルゴリズムはエージェントが提供、評価したオブジェクトが、コミュニティの方向性に合致しているかどうかで貢献度を

計算した。そこで貢献度を混乱させる攻撃に対する、EigenRumor アルゴリズムの攻撃耐性を検証する。

Dellarocas によれば[8]、オンライン取引コミュニティにおいて取引後の評価時に評価者(買い手)は、受け取った物の本来の品質より不正に高評価するか不正に低評価することにより評判システムを混乱させることを指摘している。本コミュニティモデルでは、評価値は[0,1]の範囲を取り、0 は評価なしと同一なので、不正な低評価により EigenRumor アルゴリズムを混乱させることはできない。

複数の参加者が結託して評判を混乱(攻撃)させることが考えられる。また、ネット・コミュニティの特性として1人が多数IDを使用する Sybil 攻撃も考えられるが、これは、前記結託による攻撃に一般化できる。

そこで、本稿では結託したユーザが不当に高評価する「さくら」評価を行うことに対する、EigenRumor で計算された評判スコアの攻撃耐性を検証する。

EigenRumor では、3.1 で示したように、コミュニティの種類によって e_{ij} の与え方を決定することができる。本稿で評価するモデルでは、評価する/評価しないの2択とし、1 エージェントが行う e_{ij} の合計が1となるようにする。すなわち、次式のように補正した E_{fair} を E の代わりとする。

$$\begin{aligned} E_{fair} &= [e'_{ij}], \quad (8) \\ e'_{ij} &= \frac{1}{N_i} e_{ij} \end{aligned}$$

ここで、 N_i は i の評価数である。このようなコミュニティとは、例えば、一定期間に提供された情報について、情報の評価を集約することにより情報の評判を求める場合で、参加者が使える評価ポイントを均一にし、一定期間に受けた評価を元に情報の評判を決めるような、評判システムが挙げられる。

攻撃耐性の検証は、不正オブジェクトの提供と不正評価を含んだ P 、 E を生成し、Octave[9] で実装した EigenRumor アルゴリズムに入力するというシミュレーションにより行った。

非常に活発なネット・コミュニティである、商品評価サイト@cosme[4]の参加者が行うクチコミ(評価)数、商品の被クチコミ(被評価)数の分布が共にべき分布(スケールフリー[10])であることが、筆者らが@cosme の許諾を得て分析したところ分かった。そこで、スケールフリーとなるようなリンクをプログラムで自動生成し、不正の行われていない P, E とした。ただし、各エージェント、各オブジェクトにはランダムに fitness[11]を与えて、それぞれ、評価を行う積極性、評価を受ける良さを考慮した。

5 検証結果

結託した不正エージェントの目的は、より少ない不正評価により評判システムを混乱させ、より多くのオブジェクトの評判をより高くすることであろう。今回は、不正エージェントの提供した不正オブジェクトは「不正」であることが明白で、公正なエージェントからの評価を受けないと仮定した。また、不正エージェントは攻撃の手間を最小限にするために、不正オブジェクトのみを評価し、公正なオブジェクトは評価しないとす。

シミュレーションで生成する評価リンクは、公正な評価リンク数を 1000 とし、これに加える不正評価リンク数を変化させることにより、評価リンクに占める不正評価リンクの割合を変化させた。不正オブジェクトは同じエージェントが提供するほうが提供スコアが高くなるため、1 不正エージェントにより提供した。また、不正エージェントは不正オブジェクト全てに不正評価を行うこととし、その不正エージェント数と不正オブジェクト数の割合を 1:10, 1:1, 10:1, 20:1 の 4 パターンとし、それぞれ不正評価の割合(不正率)を変化させた。1 エージェントの評価値合計は 1 となることから、不正評価数を一定とすると、結託したエージェント数を増やすことが有効な攻撃となることが考えられる。

EigenRumor が計算した評判スコアと、次のアルゴリズムで計算した評判を比較した。

EigenRumor の評判スコア

オブジェクトの被評価リンク数

オブジェクトの被評価値合計($\sum_{i \in \text{eval}(j)} e'_{ij}$)

各アルゴリズムによる値の直接比較は行えないため、不正オブジェクトより各値が追い越された正当オブジェクトの割合で比較した。なお、値はそれぞれ 5 回行ったシミュレーション結果の平均値である(図 3)

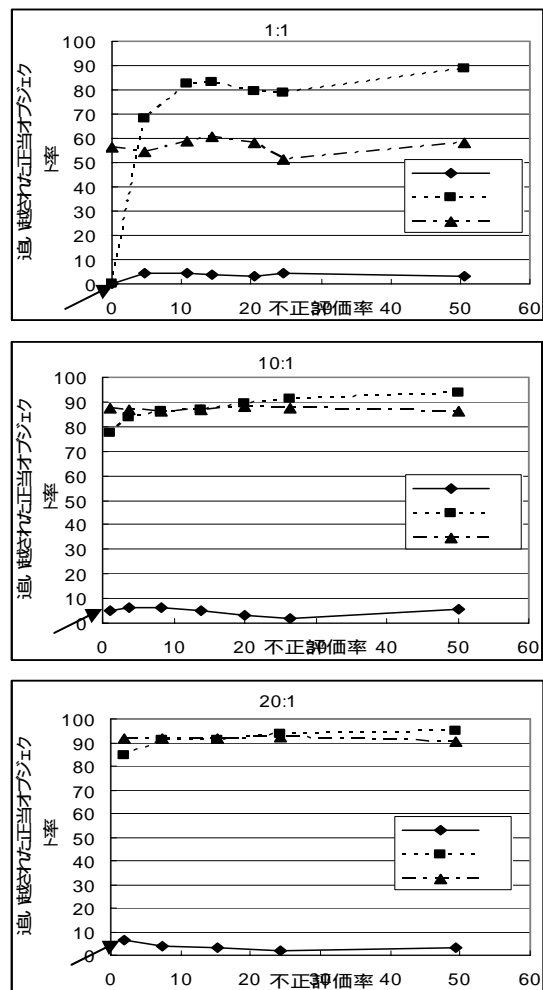
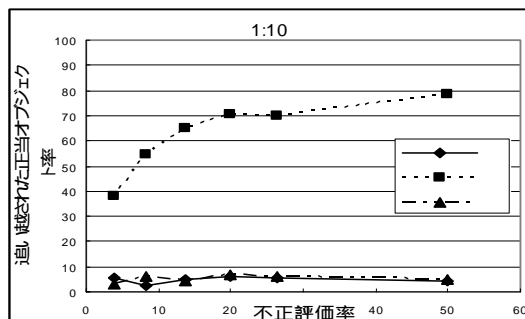


図 3 不正オブジェクトによる攻撃耐性 (単位: %)

各アルゴリズムで計算したスコアの分布は線形ではないため注意が必要であるが、次のようなことが分かる。

- 不正率が増加すると は大きな影響を受ける。
- 不正率が増加しても、 の影響は小さい。これは、1 エージェントの合計評価値が 1 であるためと考えられる。
- 1:10 20:1 と不正エージェントの割合が増加すると、 は影響が大きくなる。
- 1:10 20:1 と不正エージェントの割合が増加しても、 は影響がない。ただし、各不正エージェントの投稿数が 1 の場合はエージェント数の増加に伴って、影響が大きくなる。(図 3 の矢印)

EigenRumor による評判スコアを攻撃するためには、不正評価を行うエージェント数を増加させることが有効であることが分かったが、これについては、参加者へのコミュニティへの参加コストを上げることによる影響軽減が対策として挙げられる。

6 議論

EigenRumor アルゴリズムは、PageRank や HITS と同様に固有ベクトルを求めることを基本としている。PageRank や HITS はオブジェクト間のリンクから Authority や Hub のスコアを計算するが、EigenRumor はエージェントとオブジェクトの両方のスコアを計算し、Authority(提供)スコアと Hub(評価)スコアはエージェントの属性、Reputation はオブジェクトの属性とした(図 4)。

本稿で評価したコミュニティは一定期間の情報提供、情報評価を集めることにより評判を計算するものであった。しかし、例えば、長期間に渡って情報提供、情報評価を集める場合、新しく情報提供・情報評価を行うことに対して参加者にインセンティブを与える必要

があるが、提案アルゴリズムでは情報評価数が多いほど3つのスコアが高くなるため、一般的に古い参加者が優位になってしまう。そこで、情報評価リンクを時間により減衰させることが考えられる。例えば、 E の代わりに次の $E_{fair}^{(t)}$ が適用できる。

$$E_{fair}^{(t)} = [e_{ij}^{(t)}]$$

$$e_{ij}^{(t)} = \frac{\rho^{t-time(e_{ij})} e_{ij}}{\sum_{j \in eval(i)} \rho^{t-time(e_{ij})} e_{ij}} \quad (9)$$

ここで、 t は現在時刻、 $time(e_{ij})$ は情報評価 e_{ij} が行われた時刻、 ρ は[0,1]の値をとる減衰パラメータ。

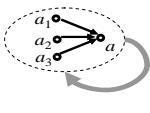
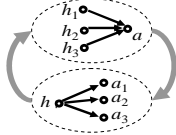
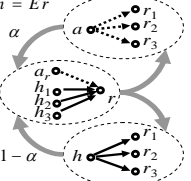
	PageRank	HITS	EigenRumor
Entities	Web page	Web page	Agent/Object
Link types	Evaluation (E)	Evaluation (E)	Evaluation (E) Provisioning (P)
Scores	Authority (\bar{a})	Authority(\bar{a}) Hub(\bar{h})	Authority(\bar{a}) } Agent Hub(\bar{h}) } Reputation(\bar{r}) Object
Algorithm	$\bar{a} = (\frac{d}{N} \mathbf{1}_N + (1-d)E^T) \bar{a}$ 	$\bar{h} = E \bar{a}$ $\bar{a} = E^T \bar{h}$ 	$\bar{r} = \alpha P^T \bar{a} + (1-\alpha) E^T \bar{h}$ $\bar{a} = P \bar{r}$ $\bar{h} = E \bar{r}$ 

図 4 PageRank, HITS とのアルゴリズム比較

7 まとめ

ネット・コミュニティにおいて、参加者の情報提供と情報評価の2つの行為を参加者から情報へのリンクとみなし、そのリンクを分析することにより、参加者のコミュニティへの貢献度(情報提供スコア、情報評価スコア)と情報の評判(評判スコア)を計算するモデルとアルゴリズム EigenRumor を示した。さらに、1 エージェントが行う評価の、評価値の合計が1となるコミュニティに EigenRumor を適用し、シミュレーションによりさくら評価からの攻撃耐性を検証し、現在主に使われている評価数や評価値の合計よりも攻撃耐性があることを明らかにした。

今後の課題として、提案アルゴリズムに対するより複雑な攻撃からの耐性評価や、負の評価投入の考慮が必要である。また、さらに別のコミュニティへの適用を検討する予定である。

参考文献

[1] S. Brin and L. Page, "The Anatomy of a

Large-scale Hypertextual Web Search Engine," In Proceedings of 7th International World Wide Web Conference, 1998.

- [2] J. M. Kleinberg, "Authoritative sources in hyperlinked environment," Journal of the ACM, Vol.46, No.5, 1999.
- [3] OKWeb, <http://okweb.jp/>
- [4] @cosme, <http://cosme.net/>
- [5] Epinions.com, <http://www.epinions.com/>
- [6] たのみこむ, <http://www.tanomi.com/>
- [7] Slashdot, <http://slashdot.jp/>
- [8] C. Dellarocas, "Immunizing Online Reputation Reporting Systems Against Unfair Ratings and Discriminatory Behavior", ACM EC00, 2000.
- [9] Octave, <http://www.octave.org/>
- [10] A. L. Barabasi, R. Albert, "Emergence of Scaling in Random Networks", SCIENCE Vol.286, 1999.
- [11] G. Bianconi, A. L. Barabasi, "Competition and multiscaling in evolving networks", Europhys. Lett. 54, 2001.