

マルチエージェント型拡張SIRモデルを用いた 情報拡散シミュレーションの評価

池田 圭佑^{1,a)} 岡田 佳之³ 榎 剛史² 鳥海 不二夫² 篠田 孝祐⁴ 風間 一洋⁵ 野田 五十樹⁶
諏訪 博彦¹ 栗原 聡^{1,b)}

概要：

2011年3月11日に発生した東日本大震災後、Twitter等ソーシャルメディアの果たした役割はとて大きく、今後起こるであろう各種災害においても重要な役割を担うことが予想される。しかし、誤った情報（流言・デマ）が広がったことも事実であり、大きな社会問題となった。災害時には、必要な情報をいち早く拡散させるだけでなく、誤った情報の拡散を早期に収束させることが重要である。これまで、我々は情報拡散モデルとして、感染症の拡散モデルとして有名なSIRモデルを拡張させたモデルを提案しているが、この手法では情報を伝搬させるかどうかを確率的状態遷移モデルで表現している。しかし、実際には人は情報に対する興味度や、情報発信元の信頼、また日常生活パターン等に基づいて情報を伝搬させるかどうかの判断を決定していると考えられる。本研究では、マルチエージェント型拡張SIRモデルを提案すると共に、それを用いた情報拡散シミュレーションの定量的な評価手法を考案し、評価を行った。

1. はじめに

ソーシャルメディアの登場により、従来までの情報流通の仕組みや人と人との繋がり・絆といった関係性の概念が大きく変わった[1]。例えば、Twitterは友人と手軽にやり取りするなど身近なコミュニケーションツールとしての役割を果たしている。さらにTwitterでの交友関係を表すフォロー・フォロワーネットワークは、現実の友人だけでなく、現実に顔を合わせたことのない人同士のコミュニケーション（著名人と一般人の直接的なやり取り等）を可能としている。

また2011年3月11日に発生した東日本大震災の際に、ソーシャルメディア、特にTwitterが避難や救援要請のための重要な情報源の一つとしてとして用いられ、自治体や

テレビ局なども積極的にTwitterを通じた情報提供を行っている[2]。Twitterは、今後起こりうる各種災害時にも被災者への有用な情報源となることが予想されている。しかし、Twitterが身近で重要な情報源になる場合には、メリットのみではなくデメリットも存在しており、その一つがデマ情報の拡散である（本稿では、白井らの定義を用い、デマを「根拠が無く、後に誤りを指摘する内容の情報が発表された情報」とする[3]）。例えば、震災時に原子力発電所も被害を受けてしまい放射能汚染に対する深刻な問題の発生や電力不足などが懸念された。この時、Twitter上では、放射性物質を取り込まないためにうがい薬や昆布などを摂取すると良いという情報や、関東地域において不足する電力は他地域における節電により補うことができるという情報がまことしやかに広まってしまった。これらの情報は後にデマであることがわかり、それらの情報がデマであり信用してはならないという訂正情報がTwitter上に流された。このように大規模な災害の場合には、被災者らは情報の真偽を確認する術がないことが予想され、デマによって深刻な被害が出てしまう恐れがある。デマや訂正情報の情報伝播メカニズムを理解することは、それらの被害を抑制するために重要である。

本稿では、ユーザをデマ情報に対する被曝回数や趣味嗜好の概念を持つエージェントとして定義することで白井らが提案したTwitter情報拡散モデルを拡張したマルチ

¹ 電気通信大学
The University of Electro-Communications
² 東京大学
The University of Tokyo
³ 大阪大学
Osaka University
⁴ 慶應義塾大学
Keio University
⁵ 和歌山大学
Wakayama University
⁶ 産業技術総合研究所
National Institute of Advanced Industrial Science and Technology
a) k-ikeda@ni.is.uec.ac.jp
b) kuri@ni.is.uec.ac.jp

エージェント型拡張 SIR モデルを用い、現実のデマ情報拡散の再現及び検証を行う。その後、シミュレーション結果を提案する評価手法を用いた評価を行う。

2章では関連研究を紹介し、3章では従来手法を整理し問題点を指摘する。4章では、その問題点を改善するための手法を提案し、5章で提案手法の検証方法を提示し、6章で実際にモデルの有効性について検証する。最後に7章で、まとめと今後の課題について述べる。

2. 関連研究

近年、Twitter に関する研究は盛んに行われている。白井らは、病気の感染モデルとして知られている SIR モデルを情報拡散モデルとして拡張し研究を行っている。このモデルは、デマ情報及びデマ訂正情報を病気を媒介するウィルスとみなし、Twitter 上での情報拡散の様子のモデル化している。その後、実際にデマが拡散した時の様子とモデルを組み込んだシミュレーションとの比較・検証を行い、提案しているモデルを用いて現実のツイート拡散が再現可能としている [3]。

石原らは、震災前後のツイートを用い、情報拡散の起点となるアカウントや情報の仲介役となるアカウントを分析している。この研究では、重要なアカウントを次数中心性と媒介中心性を各アカウント毎に求めることで特定している [4]。

三浦は、東日本大震災のツイート内容を分析し、震災時のコミュニケーション及びネガティブ表現増加の理由を、ストレスに対処するための行動であると共に、流言の増加の要因であると言及している。また、Twitter のユーザ毎にコミュニケーションが行われている場（他者とタイムラインが同一にならない）が異なる事により、とるべきアプローチが異なると述べている [5]。

Sakaki らは、地震などのイベント発生時に Twitter 上で情報発信が行われることに着目し、ツイートを監視しながらキーワード抽出を行うことで、リアルタイムに災害を発見・報告するシステムの開発と評価を行っている。開発したシステムを用いることにより、テレビ等の地震情報より早く地震発生をユーザに知らせることが可能としている [6]。

以上のように、災害と Twitter を扱った研究は多く行われている。災害以外の Twitter を用いた研究として、Stefan らは、選挙期間中に Twitter 上で行われた政治的コミュニケーションについて、情報発信源となるアカウントやどのような内容が含まれるのツイートがより拡散されやすいのか分析を行っている。その結果、多くのフォロワーを持つユーザが情報源となることや、感情を含むツイートの方がより拡散しやすいことを紹介している [7]。

このように Twitter による情報伝搬に関する研究は様々な角度から行われている。本稿では、白井らの手法を拡張

し、デマ情報への被曝回数やユーザ毎の趣味嗜好を考慮したモデルを用いることで、Twitter 上における情報拡散の再現性を改善することを目的とする。

3. 従来手法

3.1 従来モデルの概要

3.1.1 SIR モデル

白井らは SIR モデルをベースとした情報拡散モデルを提案している。SIR モデルは Kermack らによって提案された伝染病が広まっていく様子を記述する数理モデルの一種である [8][9]。

SIR モデルではある集団を構成する人を、まだ健康な人 (S : Susceptible), 病気感染し発病した人 (I : Infectious), 病気が完治した人または死亡した人 (R : Recovered) の 3 種類に分類する。このモデルでは、 S の人が I の人と接触した場合、感染率 $\rho_{(S \rightarrow I)}$ に従い病気に感染・発病することで、 I に変化する。 I の人は治癒速度 $\rho_{(I \rightarrow R)}$ に従い、病気を治癒し免疫を獲得するか死亡することにより、 R に変化する。時間の経過による S , I , R の人数の変化は以下の式 1 で表される。ここで時間 t における S の人数を $S(t)$, I の人数を $I(t)$, R の人数を $R(t)$ とすると、

$$\begin{cases} \frac{dS(t)}{dt} = -\rho_{(S \rightarrow I)} I(t) S(t) \\ \frac{dI(t)}{dt} = \rho_{(S \rightarrow I)} I(t) S(t) - \rho_{(I \rightarrow R)} I(t) \\ \frac{dR(t)}{dt} = \rho_{(I \rightarrow R)} I(t) \end{cases} \quad (1)$$

となる。また集団内の人数を N とすると、 $N = S(t) + I(t) + R(t)$ となり、常に一定である。

3.1.2 文献 [3] における情報拡散モデル

白井らは、この SIR モデルを情報拡散モデルとして拡張している。デマも人のネットワークの中で広まるという点では病気と同様であると考えられるが、デマと病気の拡散にはいくつかの違いがある。まず、病気のから回復という点である。実際の病気であれば、時間経過とともに自然治癒または死亡ということで状態が R に変化するが、デマの場合は他者から訂正情報を受け取らない限り状態が R に変化しないという点である。次に、情報拡散の場合、 $S \cdot I \cdot R$ という 3 状態のみでは不十分だということである。これはデマ情報（あるいはデマ訂正情報の場合も）を知覚した時、その情報を更に広めるかそれともそのままにしておくかで状態が違ってくるということである。そこで本来の 3 種類の状態に、新たに情報を保持しているだけの 2 種類の状態を加えた以下の 5 種類の状態を定義している。

- S : デマ情報、訂正情報の両方を見たことがない状態。
- I_{get} : デマ情報のみを見たことがある状態。訂正情報はまだ見ていない。
- I : デマ情報を投稿した状態。訂正情報はまだ見ていない。

- R_{get} : 訂正情報を見たことがある状態.
- R : 訂正情報を投稿した状態.

$$\left\{ \begin{array}{l} \frac{dS(t)}{dt} = -\frac{F}{N}I(t)S(t) \\ \frac{dI_{get}(t)}{dt} = (1 - \rho_{(S \rightarrow I)})\frac{F}{N}I(t)S(t) \\ \quad - \rho_{(I_{get} \rightarrow I)}\frac{F}{N}I_{get}(t)I(t) \\ \\ \frac{dI(t)}{dt} = \rho_{(S \rightarrow I)}\frac{F}{N}I(t)S(t) \\ \quad + \rho_{(I_{get} \rightarrow I)}\frac{F}{N}I_{get}(t)I(t) \end{array} \right. \quad (2)$$

$$\left\{ \begin{array}{l} \frac{dS(t)}{dt} = -\frac{F}{N}I(t)S(t) - \frac{F}{N}R(t)S(t) \\ \frac{dI_{get}(t)}{dt} = (1 - \rho_{(S \rightarrow I)})\frac{F}{N}I(t)S(t) \\ \quad - \rho_{(I_{get} \rightarrow I)}\frac{F}{N}I_{get}(t)I(t) \\ \quad - \frac{F}{N}I_{get}(t)R(t) \\ \\ \frac{dI(t)}{dt} = \rho_{(S \rightarrow I)}\frac{F}{N}I(t)S(t) \\ \quad + \rho_{(I_{get} \rightarrow I)}\frac{F}{N}I_{get}(t)I(t) \\ \quad - \frac{F}{N}I(t)R(t) \\ \\ \frac{dR_{get}(t)}{dt} = (1 - \rho_{(S \rightarrow R)})\frac{F}{N}R(t)S(t) \\ \quad + (1 - \rho_{(I_{get} \rightarrow R)})\frac{F}{N}I_{get}(t)R(t) \\ \quad + (1 - \rho_{(I \rightarrow R)})\frac{F}{N}I(t)R(t) \\ \quad - \rho_{(R_{get} \rightarrow R)}\frac{F}{N}R_{get}(t)R(t) \\ \\ \frac{dR(t)}{dt} = \rho_{(S \rightarrow R)}\frac{F}{N}R(t)S(t) \\ \quad + \rho_{(I_{get} \rightarrow R)}\frac{F}{N}I_{get}(t)R(t) \\ \quad + \rho_{(I \rightarrow R)}\frac{F}{N}I(t)R(t) \\ \quad + \rho_{(R_{get} \rightarrow R)}\frac{F}{N}R_{get}(t)R(t) \end{array} \right. \quad (3)$$

モデルを構成するために以下のような条件付けが行われている。ネットワーク上の全ユーザ数を N 、全ユーザの平均フォロワー数を F とする。さらに、ユーザが投稿した情報を、全てのフォロワーが見たと仮定する。SNS でのデマ拡散の場合も、デマ情報の感染速度は SIR モデル同様 S の人数に比例する。しかし Twitter では、あるユーザが投稿した文章は、基本的にそのユーザをフォローしている人しか見ることはない。このため、デマの感染速度は、各ユーザの I 以外の状態であるフォロワー数に比例する。時間 t の時点での、ユーザ 1 人あたりの S であるフォロワー数の期待値は $\frac{FS(t)}{N}$ となる。 I_{get} であるフォロワー数の期待値も同様に $\frac{FI_{get}(t)}{N}$ である。以上から、デマ訂正情報が投稿されていない状況では、感染状態は式 (2) のように記述される。なお、各感染状態における感染率を $\rho_{(S \rightarrow I)}$ (S の

ユーザがデマ情報を見たときに、 I となる確率)、 $\rho_{(I_{get} \rightarrow I)}$ 、 $\rho_{(S \rightarrow R)}$ 、 $\rho_{(I_{get} \rightarrow R)}$ 、 $\rho_{(I \rightarrow R)}$ 、 $\rho_{(R_{get} \rightarrow R)}$ とする。またデマ訂正情報をすでに見た R_{get} や R のユーザは、 S 、 I_{get} 、 I に変化しないとしている。

次に、デマ情報と訂正情報の両方が拡散している場合は、式 (3) のように記述される。

3.2 従来モデルの限界

前述した文献 [3] においては、改良すべき点をいくつか挙げる事ができる。まず、 S から I 、 I から R といった状態遷移は確率的に決まっており、これは実際の人間について考えた場合、全ユーザが同じ趣味嗜好を持っていることになり、ユーザ毎の違いを再現できていない。そのため、ユーザ毎の趣味嗜好を考慮してやる必要があると考えられる。

次に、一度デマ情報あるいはデマ訂正情報を受け取ってしまった場合、もし状態変化しなければそれ以降何度デマ情報やデマ訂正情報を受け取っても状態が変わらないという点である。実際には、一度情報を受け取るだけではデマの拡散に寄与しなかった場合でも、複数回情報を受け取ることで情報を信じてしまうことがあると考えられるため、各ユーザの情報への被曝回数数を考慮する必要がある。

最後に、実際のデマ情報の拡散についてみると、デマ情報の拡散が複数回起こった事例がいくつか確認されている。例えば、福島第一原子力発電所の事故により、東京電力管轄において電力不足が懸念される事態となった。この際に流れたのが、「関東地区に電力の融通を行うため、他の地域でも節電をするのがよい」といった内容のデマであった。このデマ情報の拡散は 3 回確認されており、それぞれ別の日に別のユーザーの間で情報が流れた。このように、実際のデマ情報はネットワーク上に一度で拡がるとは限らず、複数回に渡り拡がる事がわかっているが、この点についても考慮する必要があると言える。

本稿では、これら 3 つの改良点について、新たなモデルの提案・検証を行う。

4. マルチエージェント型拡張 SIR モデル

本稿では白井らのモデルをベースとし、エージェントの多様性と情報価値の変化を考慮した「マルチエージェント型拡張 SIR モデル」を提案する。

4.1 情報拡散のパラメータ

口コミの伝播についてモデル化した遠藤ら [10] の口コミモデルでは、情報源の信頼性及び情報の価値が重要な要素であり、その情報を信じるかどうかは受け手が持つ知識や経験により判断されると述べられている。ここで、情報の価値とは、情報の鮮度(新しさ)や情報を受取ったユーザの趣味趣向にあっているかによって評価されるものであ

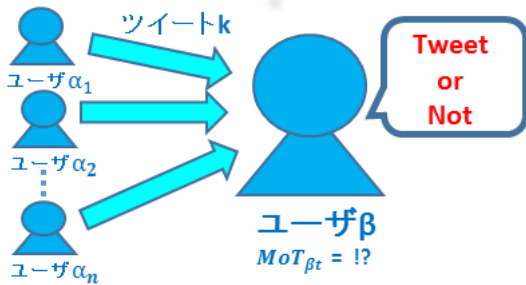


図 1 提案モデルのイメージ

る。本モデルでは、これらを考慮した情報拡散の要素となる新たなパラメータを定義する。

影響度: a

影響度 a は、情報源となるユーザが、どの程度の他者に影響度を与えるかを表すパラメータである。実際の例として、一般人よりも著名人（芸能人、政治家等）の方が信頼されやすく影響を与えやすいと考えられる。また同時にこれら著名人ユーザは情報を仲介するハブユーザと見なすことが可能である。本稿では、この値をインターネットにおけるウェブページの重要度を表す PageRank アルゴリズムを用いて定義する。これにより、フォロー・フォロワー数が多いハブユーザが強い影響を与えることを表せ、影響力の強いユーザほど値が大きくなる。

興味度: i

興味度 i は、情報を受取ったユーザがそのツイート内容を表すトピックスにどの程度興味を持っているかを表すパラメータである。これにより、各ユーザの趣味嗜好違いを表現することが可能となる。興味関心が強いほど値が大きくなる。

感度: s

感度 s は、情報を受取ったユーザがどれほど情報を信じやすいかを表すパラメータである。遠藤らの知見より、情報の真偽判断基準はユーザの知識と経験によるということから、ユーザ毎に考慮する必要がある。情報に感化されやすいユーザほど値が大きくなる。

4.2 提案モデル

本モデルは、前述したパラメータを基にユーザのツイートしたいという欲求を表す指標である MoT (Motivation of Tweet) を計算し、その値がしきい値（本稿では、特に拡散閾値と呼ぶこととする。）を越えるとユーザがつぶやき情報が拡散されるというものである。以下に、 MoT の計算式を式 (4) として示す。

$$MoT_{\beta t} = MoT_{\beta t-1} e^{-\lambda(FG-t)} + i_{k\beta} s_{\beta} \sum_n a_n \quad (4)$$

なお、 β は情報を受取りつぶやくかどうか迷っている

ユーザ、 α_n はユーザ β の情報元となるユーザの集合、 λ は忘却率、 t は現在の時刻、 FG は最初にデマ情報を受取った時刻を表すものとする。

本モデルのイメージを図 1 に示し、簡単な感染状態の遷移の概要を説明する。ユーザ β が、デマ情報を複数のフォローしているユーザ α_n から受取った場合を考える。式 (4) を用いて、ユーザ β 自身のツイート欲求 $MoT_{\beta t}$ を計算し、その値が拡散閾値を超えていればユーザ β 自身の感染状態は I となる。またしきい値を超えていない場合は、ユーザ β の感染状態は I_{get} となる。デマ訂正情報に関しても同様に考え、 $MoT_{\beta t}$ を計算し、その値が拡散閾値を超えていれば、感染状態は R とし、超えていなければ状態は R_{get} とする。ただし、状態が R または R_{get} の場合は、 I_{get} または I になることは無いものとする。

今回の提案モデルでは、各ユーザが複数回に渡って情報を受け取ることを可能にしている。これにより最初の情報ではつぶやかなくても、複数回情報を受け取ることで、関心の無い情報や信頼していなかった情報に関してもつぶやいてしまうということを再現可能としている。

また、式 (4) の右辺第 1 項から、時間の経過とともにツイートしたいという欲求が減少することが判る。例えば、地震が起こった直後に津波に注意を促すツイートが来た場合、そのツイートを拡散させたいという欲求が強いと考えられる。しかし、地震発生から数日後にそのツイートを見た場合では情報を広めたいという欲求は弱いと考えられ、ツイートをしない可能性があることを表している。つまり、この項は遠藤らが指摘している情報の鮮度について、提案モデルが扱っていることを示す。

このように、前節で指摘した文献 [3] のモデルの 3 つの問題点の内、ユーザの趣味嗜好の考慮とユーザの情報への被曝回数という 2 点について改善されると考えられる。更にもう一つの問題点である複数回に分かれて拡がる点についても、ユーザの趣味嗜好の考慮と時間経過によるツイートしたいという欲求の減少により再現可能だと考えられる。

これらの点について、6 章において本モデルを組み込んだシミュレータを用いた実験で検証を行う。

5. 評価

提案するモデルを使用するシミュレーションの結果と実際に出回ったデマとの比較によって評価を行う。

5.1 ユークリッド距離による類似度の計算

本シミュレータの結果として出力される結果は、各シミュレーションステップ時の各感染状態の人数である。また、実データのツイート内容を分析し、収集した時刻における各感染状態の人数を取得したものを用いる。これらのデータは生データであるため、今回は結果を 78 次元（実

験においてシミュレーションステップは 26 回であり、各回において S , I , R という 3 状態がある) の特徴ベクトルとして扱う。なお、実データからは各状態のそれぞれ 26 点のデータを抽出し、利用するものとする。また、それぞれのステップにおける各状態の人数の合計を分母として人数比を算出し、それを基に評価を行った。

シミュレーション結果と実データの比較には、ユークリッド距離を利用する。ユークリッド距離とは、普段我々が距離を算出するために使用しているものである。今回求めるべき距離は、対応するステップ及び状態の人数比の差であり、それらの合計が 0 に近いほど実データとシミュレーション結果が類似していると判断できる。なお、理論的には全く各データが全く類似していない場合には、今回提案するユークリッド距離の値は 78 となる。

具体的な計算法としては、以下のようになる。まず、 $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ 及び $\mathcal{Y} = \{y_1, y_2, \dots, y_n\}$ という 2 種類のベクトルについて考える。データはすでに人数比にされたものとする場合、求めるべきユークリッド距離 d は、

$$d = |x'_1 - y'_1| + |x'_2 - y'_2| + \dots + |x'_n - y'_n|$$

$$= \sum_{i=1}^n |x'_i - y'_i| \quad (5)$$

となる。

6. 実験・考察

提案モデルの有効性を検証するため実験を行う。

6.1 実験概要

本モデルによる再現性を検証するため、本モデルを用いて実際のデマツイート拡散の再現を行う。なお、今回取り扱ったデマは、東日本大震災直後に発生した千葉県市原市のコスモ石油の千葉製油所での火災時に流れたデマ情報である。この際、チェーンメールとして出回っていた「コスモ石油の爆発により有害物質が雲などに付着し、雨などといっしょに降る」といった内容のデマが Twitter にも投稿され、多数のユーザーに情報が拡散した。なお、本稿で扱う実際のデマ情報のデータは白井の研究で用いられたものと同様である。また本実験においては、情報を受取っただけでリツイートをしていない状態である“ I_{get} ”及び“ R_{get} ”については考慮しないものとする。理由は、収集したデータからはデマ情報あるいはデマ訂正状を拡散(リツイート)させたユーザ数は測定可能であるが、ツイートを読んだだけのユーザ数を測定することは困難であるからである。式(4)を用いて構築した本モデルを用い、シミュレーションを行う。この際、本モデルの特性を測るために、本モデルを組み込んだシミュレータを用い、複数回シミュレーションを行った結果の比較をする必要があると考えられる。

表 1 シミュレーション手順

ステップ 1: 表 2 のネットワークを読み込む。

ステップ 2: シミュレーション実行ステップ $t = 1$ のとき、無作為に 1 つのノードを選択し、感染状態を I に変更する。

ステップ 3: $t = 11$ のとき、無作為に 1 つのノードを選択し、感染状態を R に変更する。

ステップ 4: $t = 25$ のとき、シミュレーションを終了する。

表 2 ネットワークの設定

ノード数	50,000
リンク数(次数)の期待値	最大値 = 340 下限 = 10 パレート指数 = 0.5
リンクされやすさ	上限 = 15.0 下限 = 0.05 パレート指数 = 0.5

表 3 各パラメータの設定

興味度 i	0~1 の範囲のランダム値
感度 s	0~1 の範囲のランダム値
影響度 a	ノード毎の PageRank 値

これら 2 つの評価で使用した具体的なシミュレーション条件は文献 [3] を参考に同様の条件で行うものとする。以下の表 1 にシミュレーション手順を、表 2 にシミュレーションで用いるネットワークの設定を、表 3 にモデル内で用いているパラメータの設定を記す。

6.2 実験結果

前述の設定を用いシミュレーションを 100 回行った結果を下記に記す。

6.2.1 シミュレータ特性の評価

本モデルを用いたシミュレータの特性を評価するために、シミュレーション結果の各ステップ・各状態毎の平均と実際のデータとの比較結果を、図 2 に示す。まず図 2 より、状態“ S ”であるユーザの減少割合の様子や状態“ R ”の増加の様子は、実際のデマ拡散の様子を再現していることが確認できる。また、状態“ I ”のユーザーの割合は実データよりも少ないものの、状態遷移の様子としては実データと同様の動き方をしていると考えられる。次に、今回実データとシミュレーション結果の類似度を判定する指標となるユークリッド距離について述べる。今回実施した 100 回のシミュレーション結果それぞれのユークリッド距離を算出した所、最小値が 3.47, 最大値は 36.25 であり、それらの平均値は 15.84 であった。ここで、図 2 のユークリッド距離は 5.18 であった。また、文献 [4] の設定より得られる結果は図 3 のようになっており、この場合のユークリッド

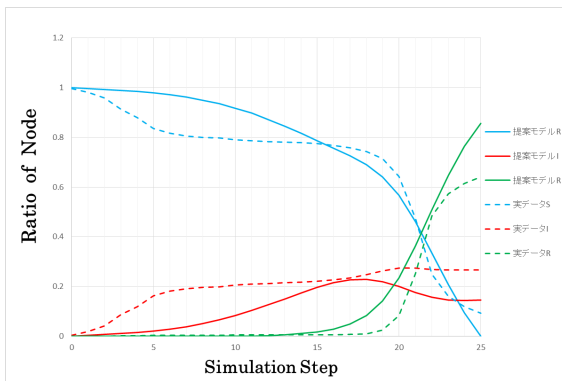


図 2 提案モデルを用いたシミュレーション結果の平均と実データとの比較

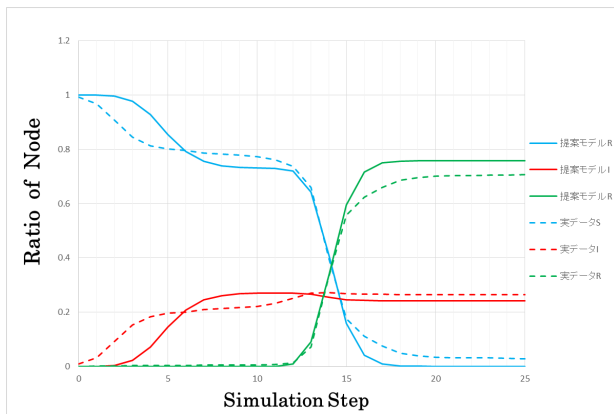


図 3 文献 [4] の設定から得られる結果

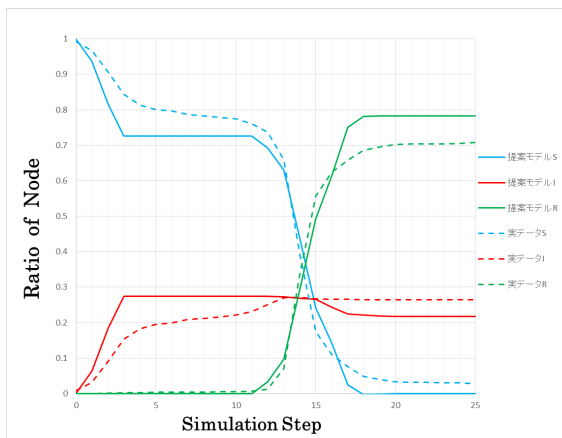


図 4 シミュレーション結果の一部と実データの比較

距離は 2.79 であり、本モデルと近い値となった。これらのことより、本モデルにおいても、現実のデマ拡散を再現できる能力を持つ可能性があることを示すことができた。

6.2.2 今後の展開

現実に関起ったデマの拡散という現象は、シミュレーションにおけるある 1 回の試行と同じであると考えられる。よって、今回行った 100 回の試行の中から、最もユークリッド距離が小さいものを図 4 に示す。また、この結果のユークリッド距離は前項で述べた通り 3.47 であった。

このことは、限りなく一致する再現が可能なパラメータや拡散ネットワークが存在する可能性を示すものである。そこで、今後の展開として、考えられるパラメータや複数の異なる拡散ネットワーク組合せに対し、シミュレーションを実行し、一致するパラメータセットの探索を行う計画である。また、拡散が急速に収束する事象や、急激に広まる事象も再現できることが推測されることから、その時の拡散ネットワークやパラメータの値から、意図的に拡散制御を行うための方策についての検討も行う計画である。

7. おわりに

東日本大震災において重要な情報源であった Twitter では、その有用性と共に誤った情報である流言やデマの拡散が問題となっていた。

本稿では、このようなデマ情報の早期収束のための、病気の感染モデルである SIR モデル基に行った白井ら [3] の研究を発展させたマルチエージェント型拡張 SIR モデルを提案した。情報拡散に寄与する要素として、情報に対する興味、情報の信じやすさ、情報源の信頼性といったパラメータを定義した。また、文献 [3] で設定されている情報に触れる回数は一度だけという制約を撤廃し、被暴露回数によってユーザの状態が変化するように改良した。更に、提案モデルの評価を行うため、各データ間のユークリッド距離を用いた評価手法を考案し、シミュレーション結果と実際のデマ拡散の様子との比較を行なった。これにより、提案モデルの妥当性の定量的な評価が行えるようになったことが示された。

今後の課題としては、6 章でも述べたように、考えられる設定の組合せの中から実際のデマをほぼ再現するようなパターンを探したい。また、デマ発生時のピークが複数回あるようなタイプの情報拡散も再現することや、どのノードにデマ訂正情報を渡せば良いのか、また、どのくらいの人数にデマ訂正情報を渡すのが効果的なのかという具体的なデマ情報の収束方法についても検討したいと考えている。

謝辞

本稿を執筆するに当たり、研究に対する助言、議論をして頂いた白井嵩士氏に感謝致します。

参考文献

- [1] 総務省:平成 23 年度情報通信白書, <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h23/pdf/index.html>, 2011
- [2] 吉次由美, 東日本大震災に見る大災害時のソーシャルメディアの役割: ツイッターを中心に, NHK 放送文化研究所, 放送研究と調査 61(7), 16-23, 2011
- [3] 白井嵩士, 榊 剛史, 鳥海 不二夫, 篠田 孝祐, 風間 一洋, 野田 五十樹, 沼尾 正行, 栗原 聡, Twitter ネットワークにおけるデマ拡散とデマ拡散防止モデルの推定, データ指向構成マイニングとシミュレーション研究会, 2012

- [4] 石原裕規, 諏訪博彦, 鳥海不二夫, 太田敏澄, 震災前後における中心性に基づく Twitter ネットワーク分析, 信学技報, vol. 112, no. 346, DE2012-29, pp. 87-92, 2012
- [5] 三浦麻子, 東日本大震災とオンラインコミュニケーションの社会心理学—そのときツイッターでは何が起きたか—, 電子情報通信学会誌, Vol.95 No.3, pp.219-223, 2012
- [6] Takeshi Sakaki, Makoto Okazaki, Yutaka Matsuo, "Earthquake Shakes Twitter Users:Real-time Event Detection by Social Sensors", WWW'10 Proceedings of the 19th international conference on World wide web, Pages 851-860, 2010
- [7] Stefan Stieglitz, Linh Dang-Xuan, "Political Communication and Influence through Microblogging-An Empirical Analysis of Sentiment in Twitter Messages and Retweet Behavior", 45th Hawaii International Conference on System Sciences, 2012
- [8] W. O. Kermack, A. G. McKendrick : A Contribution to the Mathematical Theory of Epidemics, *Proceedings of the Royal Society 115A*, pp.700-721, 1927.
- [9] 増田直紀, 今野紀雄 : 複雑ネットワークの科学, 産業図書, 2005.
- [10] 遠藤博人, 能登正人, 口コミモデルによる情報推薦システム, 電子情報通信学会技術研究報告.SITE, 技術と社会・倫理 103(78), pp.13-18, 2003