

口コミに着目した情報拡散モデルの提案および デマ情報拡散抑制手法の検証

池田 圭佑^{1,2,a)} 榊 剛史^{3,4} 鳥海 不二夫⁴ 栗原 聡¹

受付日 2017年8月28日, 再受付日 2017年10月16日/2017年11月17日,
採録日 2017年12月13日

概要: 東日本大震災や熊本地震において, Twitter などのソーシャルメディアが重要な情報源として利用された. 一方, デマ情報のような誤った情報の拡散も確認されており, デマ情報の抑制手法の確立は災害大国日本において急務である. しかし, デマ情報がどのように拡散するかは明らかになっておらず, そのため有効な抑制手法も確立されていない. 本稿では, これまでに提案した口コミに着目した情報拡散モデルにおいて「人の生活パターン」および「複数の情報源からの情報発信」を考慮した新たな情報拡散モデルを提案する. 本モデルを用いて, これまで再現性に課題のあった実際のデマ情報を再現し, 本モデルの妥当性を確認した. また, デマ情報の抑制手法の検討および評価もあわせて行った. その結果, デマ情報を否定する訂正情報をより多く拡散させるための手法が明らかになった.

キーワード: 情報拡散モデル, マルチエージェント, デマ情報の抑制, Twitter, 東日本大震災

Proposal of Information Diffusion Model Focusing on Word-of-mouth Propagation and Validation of Suppressing Methods

KEISUKE IKEDA^{1,2,a)} TAKESHI SAKAKI^{3,4} FUJIO TORIUMI⁴ SATOSHI KURIHARA¹

Received: August 28, 2017, Revised: October 16, 2017/November 17, 2017,
Accepted: December 13, 2017

Abstract: During the Great East Japan Earthquake and the Kumamoto Earthquake, people used social media such as Twitter as an important information source. On the other hand, misinformation such as false rumor was diffused. There are many disasters in Japan, we need methods to suppressing false rumor. However, it is not clear how false information diffuses, then an effective suppression method has not been established. In this paper, we propose a novel information diffusion model considering “life pattern” and “information dissemination from multiple information sources”. We confirmed the validity of our model by reproducing the actual false rumor that was not reproducible before. We also evaluated the method of suppressing false rumor. As a result, we revealed methods to spread more “correction information”.

Keywords: information diffusion model, multi-agent, suppressing false rumor, Twitter, East Japan Great Earthquake Disaster

¹ 電気通信大学大学院情報システム学研究科
The University of Electro-Communications, Chofu, Tokyo
182-8585, Japan
² 日本学術振興会特別研究員 (DC)
JSPS, Chiyoda, Tokyo 102-0083, Japan
³ 株式会社ホットリンク
Hotto Link Inc., Chiyoda, Tokyo 102-0071, Japan
⁴ 東京大学
The University of Tokyo, Bunkyo, Tokyo 113-8654, Japan
a) citrus.ki.apx500@gmail.com

1. はじめに

Twitter は, 人気のあるマイクロブログサービスであり, 多くのユーザが友人知人とのコミュニケーションや情報収集・発信のために利用している. また, Twitter は日常生活だけでなく, 震災などの災害時にも有用な情報源として利用された [9], [16]. 2011 年 3 月に発生した東日本大震

災時には、ライフライン情報や、家族・友人知人の安否情報、震災の規模などの情報が Twitter を通してやりとりされた [12]. また、2016 年 4 月に発生した熊本地震の際にも Twitter などのソーシャルメディアが活発に利用されており、今後も災害時における重要な情報源として利用されることが予想される。しかし、Twitter には、デマ情報^{*1}のような誤った情報も瞬く間に広まってしまうというデメリットも存在する。東日本大震災や熊本地震では複数のデマ情報の拡散が確認された。以下にデマ情報の例を記す。

- (1) 石油タンクが爆発、炎上した結果、有害物質が雨に混じって降ってくるので注意せよ（東日本大震災）
- (2) 関東地域において不足する電力を補うため、関西地区においても節電を行ったほうが良い（東日本大震災）
- (3) ライオンが逃げ出した（熊本地震）

災害時は情報が錯綜しており、被災者らは受け取った情報の正誤を確認することが大変困難である。そのため、デマ情報によるさらなる混乱や深刻な被害が発生する可能性がある。実際、例 (3) ではこのデマ情報により動物園の業務が妨げられたとして投稿者が逮捕された [7]. デマ情報の拡散は大きな社会問題であり、災害大国である日本ではデマ情報の拡散を早期に収束させる手法確立は急務である。そのためには、デマ情報がどのように拡散するかという情報拡散メカニズムを明らかにし、そのメカニズムに基づくデマ情報抑制手法を構築する必要がある。

本稿ではこのような背景のもと、東日本大震災を対象として研究を行う。岡田ら [8] は東日本大震災時のデマ情報について分析を行った。その結果、東日本大震災では大きく分けてデマ情報および訂正情報の拡散ピークが 1 度だけのシングルバースト型デマ拡散と、デマ情報および訂正情報の拡散ピークが複数回存在するマルチバースト型デマ拡散という 2 種類のデマ拡散現象が存在した。マルチバースト型デマ拡散は、同種のデマ情報の拡散が複数回起こり、そのつど訂正情報が拡散する現象である。そのため、シングルバースト型デマ拡散とはデマ情報の発生タイミングや情報源となるユーザの組み合わせなどに異なる部分が存在すると推測する。また、デマ情報の抑制手法を検討する際、マルチバースト型デマ拡散はなぜ複数回情報が拡散するのか、どのタイミングで訂正情報を拡散させると効率良く抑制させられるかといったことを考えなければならず、シングルバースト型デマ拡散・マルチバースト型デマ拡散両者を再現可能な情報拡散モデルの構築はきわめて重要である。さらに、社会心理学の分野では人から人に伝達されるうわさという現象において、複数回同じうわさが拡散することが知られている [15]. シングルバースト型デマ拡散・マルチバースト型デマ拡散両者を再現可能なモデルの構築は、本研究で取り組むデマ情報の抑制だけでなく、人

の情報伝達行動の理解にもつながると推測する。

これまで、筆者らは Twitter ユーザを趣味嗜好の概念を持つエージェントとして定義し、複数のエージェントが相互作用することで情報拡散現象を表現する情報拡散モデルである AIDM (Agent-based Information Diffusion Model) を提案・検証した [21]. 本モデルは、東日本大震災時に拡散が確認されたシングルバースト型デマ拡散の再現性を有する。一方で、マルチバースト型デマ拡散の再現において課題があり、情報拡散メカニズムの同定には至っていない。本稿では、実データ分析より得られた「人の生活パターン」・「複数の情報源からの情報発信」という 2 つの知見を導入し、より精緻な情報拡散モデルを構築する。あわせて、災害時にも利用可能なデマ情報の抑制手法について検証を行う。

本稿の構成について述べる。2 章では関連する研究を紹介し、本稿の位置づけを明らかにする。3 章では本稿で取り扱うデマ情報を紹介する。4 章でこれまでに提案した情報拡散モデルの特徴を述べる。5 章で現在の情報拡散モデルが持つ課題とそれらを解決する新たな情報拡散モデルを提案し、6 章で妥当性を評価する。7 章では、デマ情報抑制手法について述べ、それらの有効性を検証する。最後に 8 章でまとめを述べる。

2. 関連研究

本研究に関連する研究を紹介する。まず、情報拡散のモデル化に関する研究として、以下の研究があげられる。Takeuchi ら [6] は、コンピュータネットワーク上において、人が情報をフィルタリングしているということを考慮した情報拡散モデルを提案した。このモデルは、情報を拡散させるかの判断は、ユーザの持つ情報に対する価値によって決定される。また、情報の持つ価値にはどのようなルートで情報を得たかも含まれると述べられている。Takeuchi らの研究では、現実の人どうして構成された小規模な友人ネットワークを用いて、検証実験を行った。実験設定として、複数の情報源からの情報発信は可能であった。しかし、本研究と比べて被験者の数は 22 人ときわめて小規模であり、目的も人が情報のフィルタリングを行うということを検証するためのものであった。そのため、複数情報源からの情報発信が大規模な拡散現象となるかに対する検証はなされていない。

Serrano ら [4] は、デマ情報をつぶやいたユーザ、デマ情報であると知っているユーザ、デマ情報を否定する情報を拡散したユーザを設定し、それらのユーザの状態遷移を遷移確率を用いて遷移させることにより情報拡散現象を表現した。本モデルでは Twitter 外部からの情報の流入を考慮し、情報拡散現象を表現している。しかし、1 度デマ情報をつぶやいたユーザはデマ情報を否定するツイートをしないとしてモデル化した。ユーザごとに多様性を持たせた

*1 筆者らはデマ情報を「根拠がなく、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。

モデル作りは行われていない。また、初期値として、複数ユーザからの情報発信を行うことは可能であるが、本研究のように時間ごとに新たな情報発信者が増えていくことは考慮されていない。

小松ら [13] は、人の生活パターンを考慮した情報拡散モデルを提案した。このモデルでは、人の生活パターンとして睡眠時間を考慮し、情報を取得する時間に偏りを持たせていた。しかし、状態遷移に関しては1人ひとりのユーザの違いに着目したものではなかった。また、妥当性の検証も実際の情報拡散現象との比較は行われていない。

このように情報拡散のモデル化では、本稿で提案するモデルの各特徴を持ったモデルが多数存在するが、それら複数の特徴をあわせ持つモデル化は行われていない。また、本研究で再現に取り組んでいるマルチバースト現象に関してもこれらの研究では取り組まれていない。本稿では、複数の特徴を持つ普遍的な情報拡散モデルを構築する。

デマ情報抑制に関する研究として、様々な角度から研究が進められている。デマ情報抑制方法として、デマ情報を訂正する情報をより多くの人に伝えるという方法があげられる。このような情報を効果的に拡散させるための研究課題は「影響最大化問題」と呼ばれる。影響最大化問題はNP困難な問題であることが知られており、近似解を高速に求める研究がなされている [17], [18]。しかし、災害時のような緊急時に実ネットワークを取得し、影響力の強いユーザを特定するのは現実的ではない。そのため、災害時であっても容易に実行できる手法が必要である。

また、デマ情報の発信自体を控えさせる手法も研究されている。宮部ら [10] は、流言の拡散を防ぐための流言情報クラウドを提案した。本システムは流言情報を収集し、ユーザがつぶやこうとしている情報がこのシステム上に登録されていればその旨をユーザに伝えることにより流言の拡散を防止しようとするものである。しかし、多くのユーザに使用してもらうためには本システムの存在を十分に周知させ、ユーザに導入してもらう必要がある。実運用までのハードルが高いと考えられる。本稿で用いる抑制手法は、多くのユーザに本手法を導入してもらうためのコストがかからない。

3. ツイートデータおよびデマ情報の抽出

本章では、次章以降で述べる情報拡散モデルの構築および妥当性検証を行うため、実際に拡散したデマ情報と訂正情報の拡散の特徴を把握するための分析を行う。本研究では、東日本大震災時に拡散したデマ情報を対象とした研究を行う。そこで、本章ではシングルバースト型デマ拡散であるコスモ石油に関するデマ情報およびマルチバースト型デマ拡散である節電に関するデマ情報の概要と、それらの収集および抽出方法について述べる。なお、本章の内容は、文献 [3], [21] で用いた手法と同様であるため引用する。た

だし、引用しただけでは内容が不明瞭な部分があるため、適宜加筆する。

3.1 ツイートデータ

本研究で用いるツイートデータは、鳥海ら [22] によって行われた、東日本大震災前後における Twitter ネットワークの変化に関する研究で収集されたものを用いる。これらのツイートは TwitterAPI を用いて収集された。以下に、収集方法を引用する。

- (1) 当該期間までに 200 件以上ツイートを行ったユーザを列挙する。
- (2) 各ユーザについて 200 件ずつツイートを収集する。
- (3) 全ユーザの収集が終了した時点で、はじめのユーザに戻り改めて未収集のツイートを最大 200 件収集する。

これによって、対象となるユーザのツイートに関しては、おおむね網羅的に収集が可能である。しかし、収集が1周する間に 200 件以上ツイートしているヘビーユーザについては全ツイートを収集できてはいない。

収集されたツイートは、2011 年 3 月 5 日～2011 年 3 月 24 日までの間に日本語で投稿されたツイートである。また、本ツイートデータには、ユーザ名、ツイートの内容、投稿した時刻が含まれる。このうち本研究では、3 月 11 日の東日本大震災以降に投稿されたデマ情報ツイートおよび訂正情報ツイートを調べるため、3 月 11 日～3 月 24 日のツイートのデータを用いることとする。

3.2 デマ情報の抽出方法

東日本大震災直後に拡散したデマ情報については、インターネット上のいくつかの Web サイトで、デマ情報の内容とそれに対する正しい情報が掲載されている。本稿では、これらのデマ情報の中から、2つの事例を選択し、収集したツイートのデータから、それぞれの事例についてデマ情報および訂正情報を抽出する。また、本研究ではいくつかのデマ情報に関するキーワードを複数個設定し、そのキーワードを含むツイートを検索することで、デマ情報や訂正情報について言及しているツイートを抽出する。

ツイートデータからデマ情報と訂正情報を取り出す方法を表 1 に示す。表 1 の手順 1, 2 および手順 4, 5 については人手により行っているが、その他の手順は対象とするデータが膨大であるためプログラムを用いて自動的に分類を行っている。表 1 中の用語は以下のように定義している。なお、これらの用語の関係性をより理解してもらうために包含関係を図 1 に示す。

- 必須キーワード：ある特定のデマ情報に言及するために、必ず含まれているべきキーワード
- ネガティブキーワード：ある特定のデマ情報について、デマ情報を拡散させるツイートに含まれるキーワード
- ポジティブキーワード：ある特定のデマ情報につい

表 1 デマ情報ツイートと訂正情報ツイートの抽出法

Table 1 Extraction method of false rumor and corrected information.

手順 1	インターネットの Web 上に公開されているデマ情報を 1 つ選択する。
手順 2	ツイートがそのデマ情報に言及しているかを判別するため、デマ情報の内容から必須キーワードを設定する。
手順 3	必須キーワードでツイートを検索し、デマ情報に関する情報を含むと思われるツイート（候補ツイートデータ）をすべて抽出する。
手順 4	候補ツイートデータから、RT 数が多いデマ情報ツイートと訂正情報ツイートを複数選択する。
手順 5	選択したデマ情報ツイート、訂正情報ツイートから、ネガティブキーワードと、ポジティブキーワードを設定する。
手順 6	必須キーワード、ネガティブキーワード、ポジティブキーワードをすべて用いて検索を行い、候補ツイートデータから、デマ情報・訂正情報ツイートデータを抽出する。
手順 7	必須キーワード、ポジティブキーワードを用いて検索し、デマ情報・訂正情報ツイートデータから、デマ情報ツイートデータと、訂正情報ツイートデータを分離する。



図 1 表 1 中の用語の包含関係

Fig. 1 Containment relation of each term in Table 1.

て、訂正情報を拡散させるツイートに含まれるキーワード

- 候補ツイートデータ：全ツイートデータのうち、必須キーワードを含みある特定のデマ情報に言及していると考えられるツイート集合
- デマ情報・デマ訂正ツイートデータ：候補ツイートデータのうち、ネガティブキーワード、ポジティブキーワードを含むツイート集合
- デマ情報ツイートデータ：デマ情報・訂正情報ツイートのうち、ポジティブキーワードを含んでいないツイート集合
- 訂正情報ツイートデータ：デマ情報・訂正情報ツイートのうち、ポジティブキーワードを 1 つ以上含むツイート集合

3.3 対象とするデマ情報

本稿で対象とするデマ情報について述べる。今回、シングルバースト型デマ拡散として「コスモ石油に関するデマ情報」を対象とし、マルチバースト型デマ拡散として「節電に関するデマ情報」を対象とする。ここでシングルバースト型デマ拡散およびマルチバースト型デマ拡散の定義を述べる。

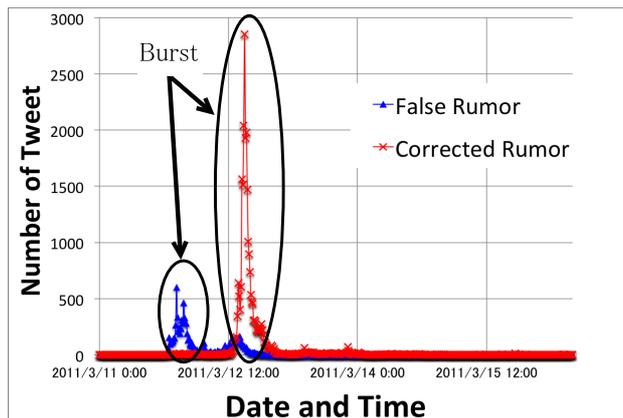


図 2 実際の拡散の様子（コスモ石油に関するデマ情報）。

文献 [21] より引用し、バースト現象がどの部分に該当するかを記載

Fig. 2 Actual diffusion process of false rumor: false rumor about Cosmo Oil Co., Ltd. Quoted from Ref. [21]. Described which part the burst phenomenon corresponds to in the figure.

- シングルバースト型デマ拡散
デマ情報や訂正情報の拡散ピークがそれぞれ 1 度だけであるデマ情報の拡散現象である。
- マルチバースト型デマ拡散
デマ情報や訂正情報の拡散ピークが複数回存在するデマ情報の拡散現象である。
具体的にデマ情報の拡散において、どの部分がバースト現象に該当するかを図 2 中に示す。

3.3.1 コスモ石油に関するデマ情報

東日本大震災直後、千葉県市原市のコスモ石油地千葉製油所で火災が発生した。この際、「コスモ石油の爆発により有害物質が雲などに付着し、雨などと一緒に降る」という内容のチェーンメールが拡散し、一部のユーザがこのメールの内容を Twitter にも投稿したことにより、Twitter 上でも拡散が起きた。その後、コスモ石油の公式ホームページに、「火災が発生したタンクに貯蔵されていたものは LP (液化天然) ガスであり、人体に及ぼす影響は非常に少ない」と発表され、デマ情報であることが分かった。

表 2 に RT 数が上位のデマ情報と訂正情報の例を記す。

表 2 コスモ石油に関するデマ情報：デマ・訂正情報の一部
Table 2 Example of Tweets: false rumor about Cosmo Oil Co., Ltd.

デマ情報	千葉市近辺に在住の方！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ので外出の際は傘かカップなどを持ち歩き、身体が雨に接触しないようにして下さい！
訂正情報	コスモ石油が否定「火災で有害物質降る」のメール連鎖 http://t.asahi.com/116s 市原市のコスモ石油千葉製油所 LPG タンクの爆発により、有害物質が雨などといっしょに飛散するという虚偽のチェーンメールが送信されています。千葉県消防地震防災課に確認したところ、そのようなことはないことを確認できました。 コスモ石油株式会社から「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」という情報について、「このような事実はありません」との発表。 http://www.cosmo-oil.co.jp/information/110312/index.html

表 3 キーワードによる各デマ情報の分類精度
Table 3 Classification accuracy of false rumor.

コスモ石油に関するデマ情報	正しく分類されたもの (件)	誤った分類をされたもの (件)	精度 (%)
デマ情報	100	0	100
訂正情報	98	2	98
節電に関するデマ情報	正しく分類されたもの (件)	誤った分類をされたもの (件)	精度 (%)
デマ情報	77	23	77
訂正情報	75	25	75

これらのツイートから選択した本デマ情報に関するツイートを抽出するために設定したキーワードを以下に記す。

- 必須キーワード：コスモ石油, 有害物質
- ネガティブキーワード：傘, カップ, レインコート
- ポジティブキーワード：デマ, ガセ, 嘘, 誤, 偽, 否定, 無害, チェーンメール, チェンメ, 事実ない, 事実ありません

ネガティブキーワードおよびポジティブキーワードは、必ず必須キーワードとともに使用され、それぞれが論理和の関係である。

この結果、全候補ツイートデータは 39,613 件、そのうちデマ情報・デマ訂正ツイートデータは 35,535 件、デマ情報ツイートデータは 9,652 件、訂正情報ツイートデータは 25,883 件であった。これらの分類が正しく行われたかを確認するため、デマ情報および訂正情報それぞれから 100 件のツイートをランダムに選択し、人手により確認した。結果を表 3 に示す。この表より、本分類は妥当であったと判断する。デマ情報ツイートデータと、訂正情報ツイートデータの実際の拡散の様子を図 2 に示す。

3.3.2 節電に関するデマ情報

東日本大震災によって発生した福島第一原子力発電所の事故により、東京電力管内の電力不足が懸念された。この際、「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情報が拡散した。特に「関西電力の社員からの情報である」といった内容のチェーンメールが広まったことにより、Twitter においても関西地区での節電が呼びかけられた。このデマ情報の場合、関東地区に電力を融通するという情報は正しいが、関東と関西では電力の周波数が異なるため、融通できる電力

には上限があり、特別な節電は必要なかった。さらに、関西電力から「当社名でお客さまに節電に関するチェーンメールを送ることはございません」との発表があった。

表 4 に RT 数が上位のデマ情報と訂正情報の例を記す。これらのツイートから選択した本デマ情報に関するツイートを抽出するために設定したキーワードを以下に記す。

- 必須キーワード：関西, 節電
- ネガティブキーワード：友達, 送電, 提供, らしい, 少し, お願い
- ポジティブキーワード：デマ, 注意, チェーンメール, チェンメ, 注意, 意味ない, 意味なし, 意味なか, 意味ありません, 必要ない, 必要なし, 必要なか, 必要ありません, 波数, Hz, 不可能, 関西電力限界, 関西電力 HP, 関西電力, 関西電力願い」に関, 関西電力サイト, 関西電力ホームページ, 関西電力ウェブ, 関西電力限界, 関西 HP, 関西, 関西電力お願い」に関, 関西電力サイト, 関西電力ホームページ, 関西電力ウェブ

ネガティブキーワードおよびポジティブキーワードは、必ず必須キーワードとともに使用され、それぞれが論理和の関係である。

この結果、全候補ツイートデータは 69,133 件、そのうちデマ情報・デマ訂正ツイートデータは 45,192 件、デマ情報ツイートデータは 15,373 件、訂正情報ツイートデータは 29,819 件であった。これらの分類が正しく行われたかを確認するため、デマ情報および訂正情報それぞれから 100 件のツイートをランダムに選択し、人手により確認した。結果を表 3 に示す。この表より、本分類は妥当であったと判断する。デマ情報ツイートデータと、訂正情報ツイートデータの実際の拡散の様子を図 3 に示す。

表 4 節電に関するデマ情報：デマ情報・訂正情報ツイートの一部
 Table 4 Example of Tweets: false rumor about power saving.

デマ情報	<p>関西電力で働いている友達からお願いなのですが、本日 18 時以降関東の電気の備蓄が底をつくらしく、中部電力や関西電力からも送電を行うらしいです。1 人が少しの節電をするだけで、関東の方が情報を得たり、病院にいる方が医療機器を使え救われます！</p> <p>関西地区にお住まいのみなさん。東北三陸沖大地震にともない、関西電力が東北電力への電力提供をはじめました。少しの節電でも立派な支援になります。電子レンジや炊飯器など、普段さしっぱなしのコンセントを今日だけでも抜き、1 人 1 人ができる節電のご協力をお願いいたします。</p>
訂正情報	<p>関西電力に聞いてみた。関西電力から東京電力への送電は行っているものの、予備電力を送電しているため、関西地域で節電しても意味が無い。また節電をお願いすることはないとのこと。忙しい中対応していただいた関西電力本店の方ありがとうございました</p> <p>不謹慎だと思ったけど関西電力に問い合わせた。周波数の関係で限界量決まっているので関係ないのだそうだ。周波数の同じ関東での節電はすべきなんだと。</p> <p>関西電力トップページにチェーンメールに関するコメントがアップ。中段以降に「今のところお客さまに特別に節電をお願いするような状況にはなく、当社で震災に関連してお客さまにチェーンメールを送ることはございませんので、ご注意ください。」</p>

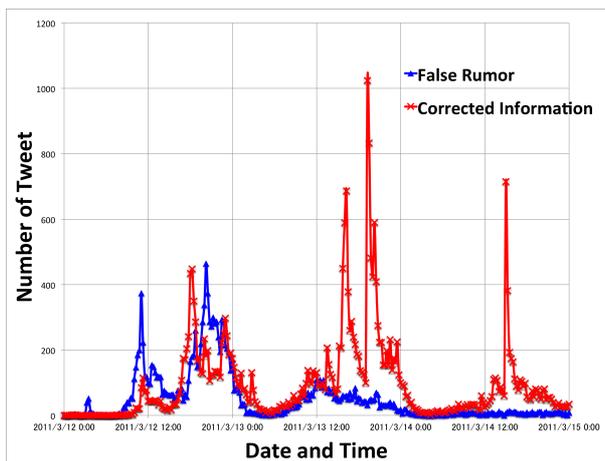


図 3 実際の拡散の様子（節電に関するデマ情報）。文献 [21] より引用

Fig. 3 Actual diffusion process of false rumor: false rumor about power saving. Quoted from Ref. [21].

3.3.3 シングルバースト型デマ拡散とマルチバースト型デマ拡散の発生に関する議論

本項では、なぜこのようにシングルバースト型デマ拡散とマルチバースト型デマ拡散のような異なる現象が発生するかを考察する。現実社会においてマルチバースト型デマ拡散となるかどうかは伝達される情報の性質によると考えられる。たとえば、コスモ石油に関するデマ情報は震災にともなう火災により、有害物質が混じった雨が降るといものである。これは火災直後に降る雨という限られた期間・地域を対象としたデマ情報であり、雨が降ってしまえばその情報の持つ意味はなくなり、その後話題となることはないと考えられ、シングルバースト型のデマ拡散となっ

たと推測する。つまり、デマ情報が示す内容が時間的・空間的にも限定されている場合は、シングルバースト型デマ拡散となる可能性が高い。また、火災時に燃えたのは LP ガスという普段料理などで日常的に使用することのある物質であり、燃焼後の物質に危険性がないことを理解しやすかったことも影響していると予想される。しかし、節電に関するデマ情報について考えてみると、節電は長期的かつ広範囲での協力が必要とされるものである。実際、関東での計画停電が多数報道 [14] されたこともあり、デマ情報の収束には日数を要し、マルチバースト型デマ拡散となったと推測する。つまり、デマ情報が示す内容が時間的な制約もなく、空間的にも広範囲を対象とする場合は、マルチバースト型デマ拡散となる可能性が高い。また、節電をすること自体は悪いことではないという情報の間違いとは関係のない要素も収束にかかる時間が伸びた理由であると推測できる。実際のツイートの中には、節電の協力自体は良いことであり、このデマ情報を否定する必要があるのかという意見も多数みられた。

以上のように、デマ情報の拡散がシングルバースト型デマ拡散となるかマルチバースト型デマ拡散となるかは情報の緊急性や継続性、内容自体の有用さなど、デマ情報の持つ意味的な要素によって決まるものと考えられる。

4. 従来型 AIDM の特徴

これまでに提案した AIDM（以下、従来型 AIDM）は、エージェントを Twitter のフォローネットワークを模した疑似的なネットワーク上に配置し、情報拡散現象をシミュレーションするモデルである [21]。従来型 AIDM は、「エージェントの多様性」、「複数回のつぶやき」、「情報経路の多

重性」という3つの特徴を持つ。本章では、これらの特徴について述べる。

4.1 エージェントの多様性

Endoら [1] の口コミ伝播の研究知見からユーザが情報を伝播させる際の重要な要素が明らかになった。この知見によると情報を拡散させる際、ユーザがその情報にどのような価値を見出すかや、情報源の信頼性が重要であるとされている。ここで「情報の価値」とは、情報の鮮度とその情報に対するユーザの興味関心の度合いである。ユーザが興味を持つ情報はユーザごとに異なり、拡散させる情報も異なると考えられる。そこで、上記のことを表現するため、次のパラメータを定義した。

影響度: a

情報源となるエージェントが持つパラメータ。本パラメータは、情報発信者が他者に与える影響を表す。一般人よりも著名人（芸能人、政治家など）の方が信頼されやすく影響を与えやすいことが予想される。

興味度: i

情報を受け取ったエージェントが受け取った情報にどの程度興味を持っているかを表すパラメータ。本パラメータにより、エージェントごとの趣味嗜好の違いを表し、実世界のような情報伝播の再現を試みる。

感度: s

情報を受け取ったエージェントがどれほど情報を信じやすいかを表すパラメータ。Endoら [1] の知見より、情報の真偽判断の基準は受け手の持つ知識と経験によるということを表す。

それぞれのエージェントは、各パラメータにより計算される「ツイートしたい」という欲求を表す MoT (Motivation of Tweet) を持ち、MoT がしきい値を超えるとエージェントがつぶやき、情報が拡散する。MoT の計算式を以下に記す。

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

ここで、 β は情報を受け取りつぶやくかどうか迷っているユーザ、 t は現在の時刻、 a_n は時刻 t においてユーザ β の情報元となるユーザの集合、 λ は忘却率、 k は受け取った情報のトピック、 FG は最初にデマ情報を受け取った時刻を表す。

4.2 複数回のつぶやき

Twitter では同一ユーザが複数回つぶやくことが可能である。人は同じトピックに対してもその情報が重要な場合や、以前つぶやいたことを忘れた場合は再びつぶやくことが考えられる。そこで、同一のデマ情報であっても複数回つぶやくことが可能なエージェントの状態遷移モデルとして ORS モデル (Outsider-Receiver-Sender モデル) を提

案・導入した。Outsider はまだデマ情報も訂正情報も知らない状態である。Receiver はデマ情報や訂正情報のどちらかあるいは両方を受け取った状態である。Sender はデマ情報や訂正情報を拡散させた状態である。1 度状態が Sender となっても、再度 Receiver に遷移することにより、新たに情報を受け取ることで再度つぶやくことが可能となる。なお、本稿では Sender のうち、デマ情報をつぶやいたユーザを“デマ情報発信者”、訂正情報をつぶやいたユーザを“訂正情報発信者”と表記する。

4.3 情報経路の多重性

Twitter では、ユーザは様々な人物をフォローしたり、フォローされたりしている。そのため、ユーザごとにタイムラインに表示される情報は異なる [11]。フォローしているユーザがいっせいに同じ内容をつぶやくことは考えにくく、タイムラインにはそのときどきで様々な情報が表示される。そのため、フォローしている人物によって各ユーザが受け取る情報は様々であり、同じ内容でも受け取るタイミングが異なると考えられる。このように、Twitter には様々な情報経路が存在するため、これを考慮する。

5. 従来型 AIDM の課題の整理と新たな情報拡散モデルの提案

本章では、従来型 AIDM の持つ課題を整理し、それらを改善する新たな情報拡散モデル（以下、新型 AIDM）を提案する。

これまでに従来型 AIDM を用いて再現実験を行った結果、従来型 AIDM がシングルバースト型デマ拡散の再現性を有することを確認した。また、マルチバースト型デマ拡散においても、デマ情報・訂正情報発信者が階段状に増加するという特徴を確認した。しかし、マルチバースト型デマ拡散のもう1つの重要な特徴である重複現象が再現できていない。マルチバースト型デマ拡散は、デマ情報または訂正情報の拡散ピークが複数回にわたるものであるが、これまでの分析により各拡散ピークにおいて重複するユーザが存在することが明らかになっている [19]。しかし、従来型 AIDM を用いた再現実験（結果の表 12）では、重複がまったく見られず、再現には至らなかった。この原因として、「人の生活パターン」および「複数情報源からの情報発信」を考慮していないことが考えられる。

5.1 人の生活パターン

Shahzadら [5] の研究により、Twitter がよく利用されている時間帯には偏りがあることが分かっている。この研究は日常生活における Twitter 利用について研究したものである。しかし、本研究が対象としているのは震災という非常事態における Twitter 利用である。そこで、震災時の Twitter 利用状況を分析した。図 4 は、2011 年 3 月 11 日

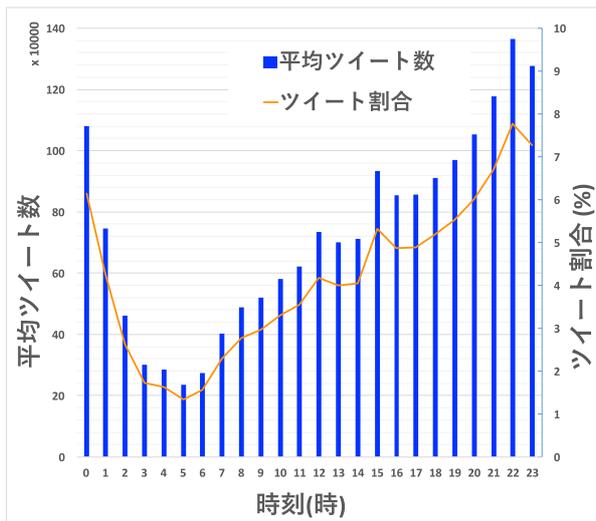


図 4 各時間における平均投稿数と投稿数の割合。
文献 [2] より引用し、日本語表記に変更

Fig. 4 Average Tweet counts and Average Tweet ratio.
Quoted from Ref. [2]. Changed to Japanese notation.

から 3 月 17 日までの 7 日間の各時間における平均ツイート数と投稿割合を示す。この図より、日中では 12 時頃や 15 時頃に投稿数が多くなっていることが分かる。これは、昼食や休憩の時間にあたり、投稿数が増えたと考えられる。また、17 時頃から再びツイート数が増えはじめ、22 時頃に 1 日の最大投稿数となっている。この時間帯は終業後からの余暇の時間帯であるためだと考えられる。23 時頃からは投稿数が減少しはじめ、早朝 5 時頃には投稿数が 1 日の最小投稿数となっている。これは、23 時から 5 時は多くのユーザが睡眠をとる時間帯であるためであると考えられる。上記より、震災時も Twitter の投稿数はユーザの生活パターンにより時間帯ごとに異なっていたことが明らかになった。

デマ情報の拡散も時間帯による変化がみられるか分析した。図 5 は、コスモ石油に関するデマ情報の時間帯ごとの投稿数である。この図から、デマ情報の投稿件数が増える時間帯は全体のツイート投稿数が増えるタイミングと一致している。また、全体的に投稿数が落ち込む深夜帯にはデマ情報の投稿数も減少していることが分かる。図 6 は、節電に関するデマ情報の投稿状況である。これもコスモ石油に関するデマ情報同様に投稿数の変化が全ツイートの投稿数に応じて変化していることが分かる。以上のことより、シングル・マルチバースト型デマ拡散の両方において、人の生活パターンにより投稿数が変化することが分かった。

そこで、本研究では時間帯ごとの Twitter 投稿割合をもとにエージェントが情報を確認するか否かを決定する。つまり、時間帯ごとに Twitter を利用するエージェント数を変化させることで、人の生活パターンを考慮する。

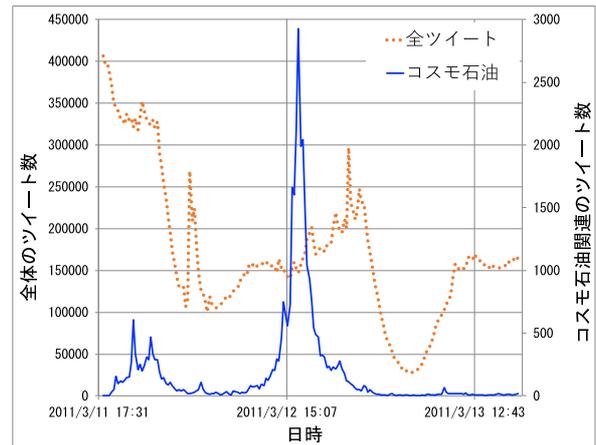


図 5 コスモ石油に関するデマ拡散とツイート投稿の関係

Fig. 5 Relationship between all tweet counts and diffusion process of false rumor: false rumor about Cosmo Oil Co., Ltd.

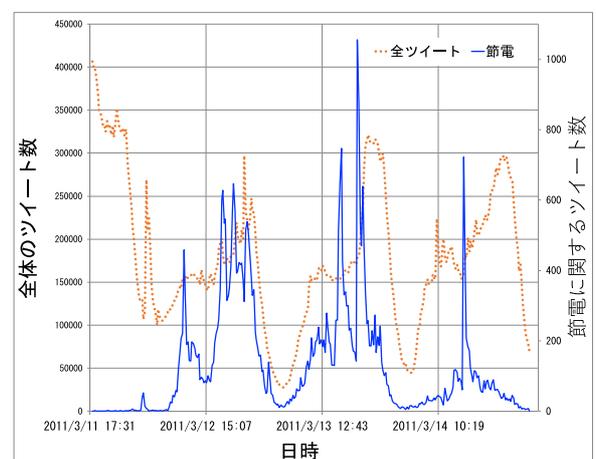


図 6 節電に関するデマ拡散とツイート投稿の関係

Fig. 6 Relationship between all tweet counts and diffusion process of false rumor: false rumor about power saving.

5.2 複数情報源からの情報発信

筆者らはこれまで Twitter を利用するユーザに着目して情報拡散モデルを構築した。その際、情報拡散はリツイート*2 (以降、RT と表記) による伝播現象であると仮定し、同じデマ情報であれば情報源となるツイートの種類は数種類程度とし、情報拡散をモデル化した。しかし、分析を行ったところ実際のデマ情報の拡散では、通常ツイート*3による拡散や情報源が複数存在することが判明した [20]。たとえば、コスモ石油に関するデマ情報において、デマ情報のソース数は 2820 件、訂正情報のソース数は 3850 件存在した。節電に関するデマ情報では、デマ情報のソース数は 6215 件、訂正情報のソース数は 8301 件存在した*4。

*2 RT (リツイート) とは、他者の投稿を引用してツイートすることで、自身のフォローにも情報を伝える方法である。

*3 通常ツイートとは、RT やリプライを用いず、単に Twitter にテキストを投稿することである。なお、Twitter で投稿できる文字数は最大 140 字という字数制限が存在する。

*4 2011 年 3 月 11 日～17 日を対象に分析。

表 5 時刻ごとの Twitter 投稿割合

Table 5 The ratio of average tweets posted at each time periods.

時刻 (時)	0	1	2	3	4	5	6	7	8	9	10	11
投稿率 (%)	6.15	4.26	2.67	1.72	1.62	1.34	1.56	2.29	2.78	2.96	3.31	3.55
時刻 (時)	12	13	14	15	16	17	18	19	20	21	22	23
投稿率 (%)	4.18	4.00	4.06	5.32	4.87	4.89	5.20	5.53	6.01	6.71	7.78	7.28

Algorithm 1 新型 AIDM におけるエージェントの振舞い

```

1: if 現在時刻における表 5 の割合に応じてエージェントがデマ情報を受け取る かつ
   同じデマを拡散していない場合 then
2:   式 (1) に従い, MoT を計算
3:   if MoT > しきい値 then
4:     状態を Sender に遷移し, そのエージェントのフォロワに
       デマ情報を拡散
5:   else
6:     状態を Receiver に遷移
7:   end if
8: end if
9: if 状態が Sender かつ 新たなデマ情報を受け取る then
10:  状態を Receiver に遷移
11: end if
    新たにデマ情報を受け取ったら, 同様に繰り返す

```

よって, 新型 AIDM では複数の情報源からデマおよび訂正情報の発信が行われるものとする. 複数情報源からの情報発信を考慮するため, 各シミュレーションステップ時に一定の条件 (ノードが持つフォロワ数が 100 以上^{*5}) を満たしたノードから無作為に選択し, 新規のデマ情報の発信源とする. その際, 人の生活パターンを考慮し, 時間帯ごとにデマ情報を発信するノード数を変化させる. なお, シミュレーションの全ステップにおけるデマ情報の最大投入数は実際のデマ情報のソース数と実際のユーザ数およびシミュレーションで使用するノード数をもとに決定する. また, 訂正情報の発信に関しても同様とする.

5.3 新型 AIDM における各エージェントの振舞い

新型 AIDM における各エージェントの振舞いを擬似コード (Algorithm 1) として示す. この擬似コードを, ユーザ β がデマ情報を受け取った場合を用いて説明する. まず, ユーザ β は現在時刻における表 5 の割合に応じてデマ情報を受け取るかを決定する. なお, 表 5 は実データ (図 4) から取得した時間帯ごとのツイート投稿割合を基に作成した. ユーザ β がデマ情報を受け取った場合は, 式 (1) に従い MoT を計算する. もし, MoT がしきい値を超えていれば, ユーザ β はデマ情報をリツイートし, デマ情報をユーザ β のフォロワに情報を伝播させる. この際, MoT を 0 にするのではなく, 4 割だけ残すこととする^{*6}. なお, この 4 割という値は, 複数の条件下での実験結果から決定し

^{*5} これは, 文献 [3] と同様の条件である.

^{*6} 人はすべてのことを忘れることはないかと仮定する.

た. なお, 具体的な決定方法などについては, 6.1 節および付録で述べる. もし, MoT がしきい値を超えていなければ, ユーザ β はそのデマ情報をリツイートしない. その後, ユーザ β が新たなデマ情報を受け取ると再度 MoT を計算し, しきい値以上であればその情報がリツイートされ, 情報が伝播する. ユーザ β が 1 度デマ情報を拡散していたとしても, 異なるデマ情報を受け取った場合であれば同様に振る舞う. 訂正情報を受け取った場合も, 同様である.

6. 新型 AIDM の妥当性検証実験

本章では, 新型 AIDM の妥当性を検証するための実験について述べる. 妥当性検証のため, 東日本大震災で拡散が確認された 2 種類のデマ情報を再現する. 対象とするデマ拡散は, 「コスモ石油に関するデマ情報」(シングルバースト型デマ拡散), 「節電に関するデマ情報」(マルチバースト型デマ拡散) である.

6.1 実験手順および評価手法

本研究では, 提案モデルの妥当性を実際のデマ情報の拡散を再現可能かどうかにより評価する. そのため, 実際にデマ情報が拡散した期間および訂正情報が投入された時間をもとにシミュレーションを行う. そこで, まず各デマ情報が発生した際の状況を述べる. 実データ分析より, コスモ石油に関するデマ拡散は 2011 年 3 月 11 日 19 時頃から拡散しはじめ, 訂正情報は同日 20 時頃拡散しはじめたことが分かった. その後, 3 月 13 日 18 時頃には拡散は落ち着き, ほぼ収まったものと見なす. 同様に, 節電に関するデマ拡散は 2011 年 3 月 11 日 18 時 45 分頃拡散しはじめ, 訂正情報も同一時刻に拡散しはじめたことが分かった. 収束はコスモ石油のデマ拡散よりも遅く, 拡散し始めてから数日後の 3 月 15 日 23 時 59 分頃には拡散が落ち着いたため, ほぼ収束したと見なす. 上記より, コスモ石油・節電に関するデマ拡散の実験手順を表 8, 表 9 にそれぞれ記す. 本実験で使用するネットワークおよび各エージェントのパラメータ設定をそれぞれ表 6 と表 7 に記す. 表 6 は, スケールフリーネットワークを生成するための条件を示しており, この条件によりノードの次数が最大 3000 となるようなベキ分布に従うネットワークを生成する. 表 7 は, MoT の計算に必要なパラメータの設定を記している. ここで, 興味度, 感度, 影響度はエージェントごとに異なる

表 6 ネットワークの設定

Table 6 Parameters of network.

ノード数	100,000
リンク数 (次数) の期待値	最大値 = 3000 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 7 各パラメータの設定

Table 7 Parameter of agents.

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	ノードごとの PageRank 値
忘却率 λ	1/8
しきい値	0.00003

表 8 シングルバースト型デマ拡散の実験手順

Table 8 Procedure of simulation: Single-burst type false rumor diffusion.

ステップ 1	表 6 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 $t = 0$ のとき、無作為に 1 つのノードを選択し、感染状態を“デマ情報発信者”に変更する その後、表 5 の割合に応じ、新たなデマ情報を投入する
ステップ 3	$t = 4$ のとき、無作為に 1 つのノードを選択し、感染状態を“訂正情報発信者”に変更する その後、表 5 の割合に応じ、新たなデマ情報を投入する
ステップ 4	$t = 188$ のとき、シミュレーションを終了する

表 9 マルチバースト型デマ拡散の実験手順

Table 9 Procedure of simulation: Multi-burst type false rumor diffusion.

ステップ 1	表 6 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 $t = 0$ のとき、無作為に 1 つのノードを選択し、感染状態を“デマ情報発信者”に変更する その後、表 5 の割合に応じ、新たなデマ情報を投入する
ステップ 3	$t = 0$ のとき、無作為に 1 つのノードを選択し、感染状態を“訂正情報発信者”に変更する その後、表 5 の割合に応じ、新たなデマ情報を投入する
ステップ 4	$t = 404$ のとき、シミュレーションを終了する

る値とすることでエージェントの多様性を表現する。忘却率としきい値および前述した「MoTを残す割合」は予備実験の結果を参考とし、決定した。本予備実験では、各パラメータを変更して複数回シミュレーションを行い、平均の情報発信者数を比較した。予備実験で再現したデマ情報は

「コスモ石油に関するデマ情報」であり、情報発信者の延べ数が数千から数万人規模になるようなパラメータを採用した。これは、実際のコスモ石油に関するデマ情報の拡散が数万人規模の拡散であったこと、平均の情報発信者数を予備実験の評価に使うことからこのような目安とした。平均を用いた理由は、すべてのシミュレーション結果を目で見てパラメータの調整をすることが困難なためである。各パラメータの変更は情報発信数の増加の速さや、最終的な情報発信数の決定に影響を及ぼしており、大きく影響を及ぼすのは、「しきい値」と「MoTを残す割合」である。「忘却率」の変更は大きく情報発信者数に影響を及ぼさなかった。ただし、これらのパラメータは実データの分析より得たものではないことに注意が必要である。なお、予備実験の方法と各パラメータの変更が及ぼす影響については付録で詳しく述べる。

人の生活パターンを考慮するため各時刻においてタイムラインを読むことのできるユーザは表 5 の割合とし、シミュレーション 1 ステップは実時間の 15 分とする。本実験での複数情報源からの情報発信の最大数は、デマ情報・訂正情報ともに 200 件までとする。また、1 ステップに投入できる情報数は 10 件までとする。これは、シミュレーションで用いるネットワークが実際のネットワークよりも小さいため、際限なく情報発信が行われるのを防ぐためである。シミュレーション回数は 5000 回とし、後述する類似度が最も小さくなるものを結果とする。これは、筆者らの「デマ情報の投稿は日常的に行われており、大規模な拡散は偶然である」という考えからであり、文献 [21] と同様の条件である。

次に、実験結果の評価方法について述べる。本稿では、次の 2 つの手法を用いて提案モデルの妥当性を評価する。**類似度**

本研究では、実データとシミュレーション結果のユークリッド距離を用いて類似性を評価する。このユークリッド距離が 0 に近いほど、シミュレーション結果が実データに類似していると見なす。各シミュレーションステップのある状態の人数比を $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ とし、実データのある状態の人数比を $\mathcal{Y} = \{y_1, y_2, \dots, y_n\}$ と表す場合、ユークリッド距離 d は、以下のように計算される。

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

重複率

マルチバースト型デマ拡散の評価では、各バースト期間の組合せにおいて重複するユーザの割合を実データと比較し、評価する。表 10 に示す期間において、各

表 10 対象とする期間の設定

Table 10 Target period.

デマ情報	実データ (2011年3月)	従来型モデル	新型 AIDM
第1期間	12日0時00分 - 12日14時59分	Step0 - 5	Step21 - 80
第2期間	12日15時00分 - 13日5時59分	Step6 - 11	Step81 - 140
第3期間	13日6時00分00秒 - 13日23時59分	Step11 - 24	Step141 - 212
訂正情報	実データ (2011年3月)	従来型 AIDM	新型 AIDM
第1期間	12日9時30分 - 13日5時59分	Step3 - 11	Step57 - 140
第2期間	13日6時00分 - 14日5時59分	Step12 - 17	Step141 - 237
第3期間	14日6時00分 - 15日23時59分	Step18 - 24	Step238 - 404

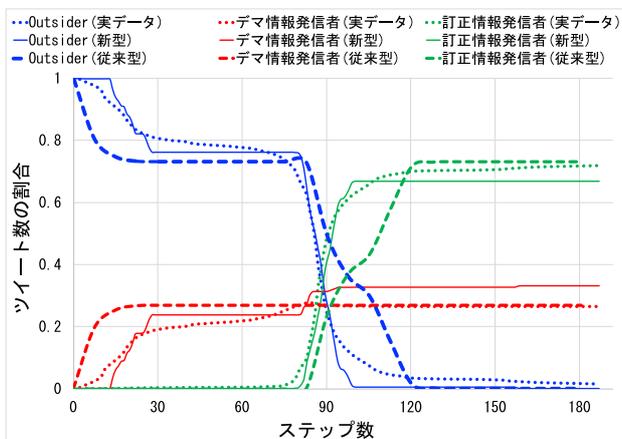


図 7 コスモ石油に関するデマ情報の再現結果

Fig. 7 Result of reproduction experiment: false rumor about Cosmo Oil Co., Ltd.

期間ごとのユーザーリストを作成し、ユーザーの重複の有無とその割合を調べる。シミュレーション結果についても同様にユーザーの重複率を分析し、実データと比較する。

また、従来型 AIDM との比較もあわせて行う。従来型 AIDM の実験設定は文献 [21] と同様である。ただし、類似度については従来型 AIDM と新型 AIDM でシミュレーションのステップ数が異なるため、各実験のステップ数で類似度を割った値を比較する。

6.2 実験結果

図 7 にコスモ石油に関するデマ情報の再現結果を示し、表 11 に 1 ステップあたりの類似度を記す。この図より、デマ情報発信者は実データよりも拡散が少し遅れており、ステップ 80 あたりで階段状に発信者数が増えているものの、おおむね実データに即していることが分かる。また、訂正情報発信者はほぼ実データに即して増加していることが分かる。類似度は、Outsider、デマ情報発信者、訂正情

表 11 類似度。比較のため、ステップ数で類似度を割った値を記す

Table 11 Similarity: describe the similarity per unit step, because the simulation steps are different.

コスモ石油	Outsider	デマ情報発信者	訂正情報発信者	平均
新型 AIDM	0.00267	0.00362	0.00241	0.00287
従来型 AIDM	0.0131	0.00918	0.00969	0.0106
節電	Outsider	デマ情報発信者	訂正情報発信者	平均
新型 AIDM	0.00400	0.00190	0.00296	0.00295
従来型 AIDM	0.0190	0.00491	0.0189	0.0143

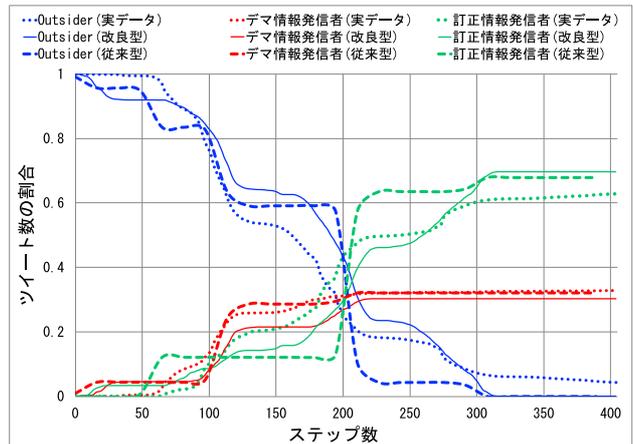


図 8 節電に関するデマ情報の再現結果

Fig. 8 Result of reproduction experiment: false rumor about power saving.

表 12 各期間の組合せにおけるユーザーの重複率

Table 12 Overlapping rate.

実データ	第1と第2期間	第1と第3期間	第3と第2期間
デマ情報 (%)	1.96	1.09	1.78
訂正情報 (%)	5.84	2.40	3.20
従来型 AIDM	第1と第2期間	第1と第3期間	第3と第2期間
デマ情報 (%)	0	0	0
訂正情報 (%)	0	0	0
新型 AIDM	第1と第2期間	第1と第3期間	第3と第2期間
デマ情報 (%)	8.39	7.39	16.54
訂正情報 (%)	22.8	21.8	35.0

報発信者ともに従来型 AIDM よりも距離が短くなっている。よって、新型 AIDM はシングルバースト型デマ拡散であるコスモ石油に関するデマ情報の再現性を有することが明らかになった。

次に、節電に関するデマ情報の再現結果を図 8 に示す。この図から、デマ情報発信者、訂正情報発信者の増加の様子は一部実データと乖離している部分もあるが、おおむね再現できている。また、マルチバースト型デマ拡散の重要な特徴である階段状に情報発信者が増える現象を再現できた。表 11 に記した類似度の比較では、訂正情報発信者の距離が従来型 AIDM よりも大きくなってしまっている。しかし、Outsider やデマ情報発信者は距離が短くなっており、平均的には類似性が増加している。各期間における重複率を表 12 に記す。この表から、新型 AIDM による再現では、各期間において同一のユーザーが投稿していることが

分かる。実データと比較すると、重複率は大きくなってしまっている。これは実際のネットワーク構造よりもシミュレーションで使用したネットワークが小さく、同一ユーザに情報が伝播しやすいことが原因として考えられる。今後、他のネットワークや実際のネットワークを使用して検証を進める予定である。しかし、従来型 AIDM では、重複がいったい発生しなかったが、「人の生活パタン」と「複数の情報源からの情報発信」を考慮することにより実際の現象と同様の現象を再現できた。よって、新型 AIDM はマルチバースト型デマ拡散である節電に関するデマ情報の再現性を有することが明らかになった。

以上の結果より、新型 AIDM は同一のモデルにもかかわらず、シングル・マルチ両方のバースト現象を再現可能であることが明らかになった。

続いて、本稿で提案した2つの要素の必要性に関して述べる。本稿では、述べていないが我々はこれまで従来型 AIDM において生活パタンのみを考慮した情報拡散モデルで再現実験を行った。しかし、このモデルではマルチバースト型デマ拡散の再現には至らなかった。今回、複数の情報源からの情報発信を考慮することによりマルチバースト型デマ拡散を再現できたことから、複数の情報源からの情報発信は必要な要素である。また、複数の情報源からの情報発信を行う際、人の生活パタンを考慮して情報発信を行うため、人の生活パタンを考慮することも新型 AIDM において重要な要素である。

また、3.3.3 項で述べたようにデマ情報の拡散がシングルバースト型デマ拡散となるかマルチバースト型デマ拡散となるかは情報の緊急性やもとなる現象の継続性、内容自体の有用さなど、デマ情報の持つ質的な要素によって変わるものと考えられる。人の生活パタンの考慮は、この情報拡散の継続性を表現するために重要であると考えられる。なぜなら、人の活動時間（今回であれば Twitter の利用時間帯）を考慮しなければ、このような数日間継続して拡散するタイプのデマ拡散には対応できないと予想される。つまり、単にどの時間帯にも偏りなく情報発信を行えば、情報発信が疎な時間帯、密な時間帯が形成されず、マルチバースト型デマ拡散の再現はできない。この点からも人の生活パタンを考慮することは重要である。

よって、本稿で提案した2つの要素両方がそろえることがマルチバースト型デマ拡散を再現するために必要な条件であると考えられる。

7. デマ情報抑制手法の検証

本章では、災害時に実際に利用してもらうためのデマ情報抑制手法の有効性を検証する。震災時の利用を想定した抑制手法に必要な要件として、「抑制手法の導入・実行に特別なコストが必要ない」ことがあげられる。震災時、政府や地方自治体、企業などは震災自体の状況把握を行い、必

要な救助活動および避難活動などを行う必要があり、混乱が予想される。そのため、デマ情報の抑制手法としては、収束させるためのコストが可能な限り低く、そして、より効果的な手法が望ましい。筆者らは、これまで訂正情報拡散の起点となるユーザに焦点を当て、あるユーザに訂正情報を投稿するように依頼するという仮定のもと3種類の抑制手法を提案した [3]。

- 抑制手法 A：デマ情報拡散の起点となったユーザに、訂正情報拡散の起点になってもらう。
- 抑制手法 B：全ユーザのうち、ネットワークのハブとなっているユーザ 1 人に、訂正情報拡散の起点になってもらう。本研究ではハブユーザを 2000 人以上のフォロワーを持つユーザとする。
- 抑制手法 C：デマ情報を投稿したユーザのうち、最もフォロワーが多いユーザに、訂正情報拡散の起点になってもらう。

しかし、これらの手法の検証はシングルバースト型デマ拡散のみで行っており、マルチバースト型デマ拡散に対しても有効であるかの検証は行っていない。本稿では、これらの手法がシングル・マルチバースト型デマ拡散の両方に有効であるかを検証する。

実験の設定および手順は、6 章と同様である。ただし、訂正情報の起点ユーザ選択では、実際のデマ情報の拡散でデマに関する注意を促す公式アナウンスがあった日時に各手法の選択基準により訂正情報の投稿を依頼するものとする。また、実データ分析では公式アナウンス以前にも訂正情報の発信が確認されているため、各エージェントの自発的な訂正情報の発信はそのまま続ける。各デマ情報の公式アナウンスのあった日時は、3 月 12 日 14 時頃（コスモ石油に関するデマ情報）と 3 月 15 日 15 時頃（節電に関するデマ情報）である。よって、シミュレーションではそれぞれ 76 ステップ目と 177 ステップ目に公式情報の投稿を依頼する。また、各手法との比較のため、各公式情報の投入タイミングで無作為に 100 人以上のフォロワーを持つユーザを選び、訂正情報をつぶやいてもらう（ランダム選択）。シミュレーションは 100 回行い、平均の訂正情報発信者の増加をみる。この試行回数設定とそれらの平均をとることは文献 [3] と同様の条件である。また、今後現実社会で発生するデマ情報の拡散は、拡散の経路や規模がまったく同じ拡散現象となるとは考えられない。そのため、抑制手法の検証では平均的に良い結果を得られる手法—より早くかつ迅速に訂正情報を拡散させられる手法—が有効であると判断する。

7.1 拡散抑制手法の有効性検証実験

図 9 にコスモ石油に関するデマ情報における各抑制手法の実験結果を記す。この図より、ランダム選択よりも各抑制手法の方が短時間により多くの訂正情報を拡散可能であ

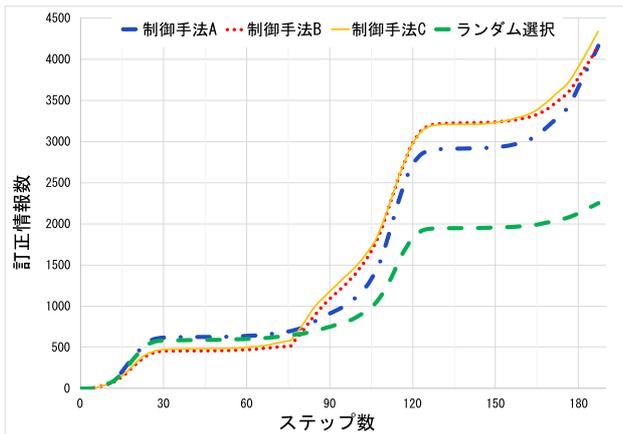


図 9 各抑制手法ごとの訂正情報発信者数 (コスモ石油に関するデマ情報)

Fig. 9 Simulation result of control methods: false rumor about Cosmo Oil Co., Ltd.

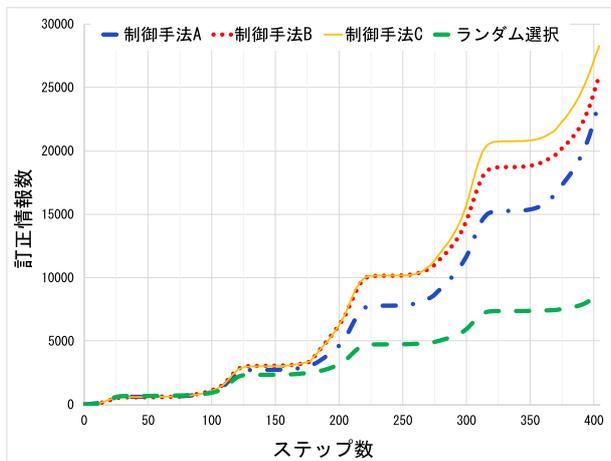


図 10 各抑制手法ごとの訂正情報発信者数 (節電に関するデマ情報)

Fig. 10 Simulation result of control methods: false rumor about power saving.

ることが分かった。また、抑制手法 A よりも抑制手法 B・C の方が多くの訂正情報拡散につながった。

次に、節電に関するデマ情報における各抑制手法の実験結果を図 10 に示す。この図より、ランダム選択よりも各抑制手法の方が素早く多くの訂正情報を拡散できた。また、抑制手法 A よりも抑制手法 B・C の方が訂正情報をより速く拡散させることができた。なお、マルチバースト型デマ拡散では抑制手法 B よりも抑制手法 C の方が優位であった。

訂正情報の増加は、多くのユーザに訂正情報を周知でき、デマ情報への警戒につながる。そのため、抑制手法 B および C はデマ情報の抑制に有効である。また、抑制手法 B は必ずしもハブユーザに訂正情報の発信を依頼していないにもかかわらず、ハブユーザに依頼した場合と同等以上の結果を得た。この結果は実際の災害での利用を想定する場合、きわめて重要である。なぜなら、ネットワーク全体の

ハブに訂正情報の拡散を依頼する際、必ずしもハブユーザがそのデマ情報を知っているとは限らず、デマ情報の詳細を説明し、訂正情報を拡散してもらうまでにコストがかかる。しかし、デマ情報を一度でもつぶやいたユーザであれば、デマ情報の内容を理解しているため、その分コストが低いと推測される。複数のデマ情報が拡散している状況を考えた場合、手法 B ではハブユーザの協力を得られない恐れもある。これは様々な情報がハブユーザに集まることで、ハブユーザの処理能力を超えてしまい、デマ情報の訂正依頼に気づいてもらえないことが推測されるためである。

8. おわりに

東日本大震災や熊本地震において、Twitter などのソーシャルメディアが有用な情報源として積極的に利用されており、今後の災害時にも利用されることが予想される。一方で、ソーシャルメディアではデマ情報の拡散が社会問題化しており、震災などの混乱した状況下ではデマ情報によりさらなる混乱につながる恐れもある。そのため、デマ情報の拡散を抑制するための手法の確立が急務である。しかし、ソーシャルメディア上で情報が拡散するメカニズムは明らかになっておらず、震災時に利用可能なデマ情報の抑制手法には課題があった。

これまで Twitter ユーザの多様性や Twitter ネットワークの持つ特徴を考慮した情報拡散モデル (従来型 AIDM) を構築したが、いくつかの課題があり、情報拡散メカニズム同定には至っていなかった。本稿では、従来型 AIDM の持つ課題を明らかにし、それらの課題を解決するため「人の生活パターン」、「複数情報源からの情報発信」を導入した新型 AIDM を提案した。本モデルを用いてシングルバースト型デマ拡散とこれまで限定的にしか再現できていなかったマルチバースト型デマ拡散を再現し、新型 AIDM の妥当性を検証した。

また、この新型モデルを用い、震災時に利用されることを想定した拡散抑制手法についても検証した。デマ情報の抑制を「訂正情報を短時間でより多くのユーザに届ける」とこと考え、実験を行った。その結果、ネットワークのハブとなっているユーザや、デマ情報をつぶやいたユーザの中で最も多くのフォロワーを持つユーザに訂正情報をつぶやいてもらうことが有効であると判明した。

今後の課題について述べる。まずは、提案した AIDM の持つ各パラメータをどのように推定するかという点である。AIDM では、影響度・興味度・感度・忘却率・しきい値といったパラメータを導入した。しかし、これらのパラメータを実データからは推定できていない。実際の影響度は各ユーザの回数などをもとに計算可能であると考えられ、その他のパラメータに関しては被験者実験により推定する方法や、Twitter ユーザの行動ログやアクセスログを解析することにより推定する方法があると考えられる。今後、このよ

うな手法をもとにパラメータ推定を行い、より現実的なシミュレーションを実施したい。そして、これらのパラメータを活かした制御手法の提案も行う予定である。

次に、新型 AIDM の普遍性を確かめるため、様々な対象・文化圏で発生する拡散現象の再現に取り組む予定である。また、デマ情報の拡散は災害時以外でも問題となっており、日常において発生するデマ情報を抑制するための方法を提案し、検証に取り組む。

謝辞 本研究は JSPS 科研費 JP16J04396 の助成を受けたものである。

参考文献

- [1] Endo, H. and Noto, M.: A word-of-mouth information recommender system considering information reliability and user preferences, *IEEE International Conference on Systems, Man and Cybernetics, 2003*, Vol.3, IEEE, pp.2990-2995 (2003).
- [2] Ikeda, K., Sakaki, T., Toriumi, F. and Kurihara, S.: "An Examination of a Novel Information Diffusion Model: Considering of Twitter User and Twitter System Features," *Autonomous Agents and Multiagent Systems AAMAS 2016 Workshops, Best Papers Singapore, Singapore, May 9-10, 2016 Revised Selected Papers*, Springer (2016).
- [3] Okada, Y., Ikeda, K., Shinoda, K., Toriumi, F., Sakaki, T., Kazama, K., Numao, M., Noda, I. and Kurihara, S.: SIR-extended information diffusion model of false rumor and its prevention strategy for Twitter, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol.18, No.4, pp.598-607 (2014).
- [4] Serrano, E., Iglesias, C.Á. and Garijo, M.: A Novel agent-based rumor spreading model in Twitter, *Proc. 24th International Conference on World Wide Web*, ACM, pp.811-814 (2015).
- [5] Shahzad, B. and Alwagait, E.: Best and the worst times to tweet: an experimental study, *WSEAS, 15th International Conference on Mathematics and Computers in Business and Economics (MCBE'14), Proc. 15th International Conference on Mathematics and Computers in Business and Economics*, pp.122-126 (2014).
- [6] Takeuchi, S., Kamahara, J., Shimojo, S. and Miyahara, H.: Human-network-based filtering: the information propagation model based on word-of-mouth communication, *Proc. 2003 Symposium on Applications and the Internet, 2003*, IEEE, pp.40-47 (2003).
- [7] 安藤健二:「ライオン逃げた」熊本地震のデマ情報を拡散した疑い 20歳男を逮捕, *The Huffington Post Japan, Ltd.* (参照 2017-8-23).
- [8] 岡田佳之, 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: 拡張 SIR モデルによる Twitter でのデマ拡散過程の解析, *人工知能学会全国大会論文集* (2013).
- [9] 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, *放送研究と調査*, Vol.61, No.7, pp.16-23 (2011).
- [10] 宮部真衣, 灘本明代, 荒牧英治ほか: 人間による訂正情報に着目した流言拡散防止サービスの構築, *情報処理学会論文誌*, Vol.55, No.1, pp.563-573 (2014).
- [11] 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こったか, *電子情報通信学会誌*, Vol.95, No.3, pp.219-223 (2012).
- [12] 執行文子: 東日本大震災・被災者はメディアをどのように利用したのか: ネットユーザーに対するオンライングループインタビュー調査から, *放送研究と調査*, Vol.61, No.9, pp.18-30 (2011).
- [13] 小松琢也, 鈴木育男, 山本雅人, 古川正志: 時間遅れを考慮した情報伝播におけるトポロジーの影響について, *精密工学会学術講演会講演論文集 2013 年度精密工学会春季大会*, 公益社団法人精密工学会, pp.285-286 (2013).
- [14] 沼田宗純, 目黒公郎: 2011 年東日本大震災のテレビ報道を用いた「計画停電・電力不足・節電」に関する課題の時系列分析, *平成 24 年度土木学会年次学術講演会* (2012).
- [15] 川上善郎: うわさが走る: 情報伝播の社会心理, *サイエンス社* (1997).
- [16] 総務省: 平成 23 年度情報通信白書 (2011).
- [17] 大原剛三, 斉藤和巳, 木村昌弘, 元田 浩: 情報拡散モデルに基づく社会ネットワークの影響度分析, *オペレーションズ・リサーチ: 経営の科学*, Vol.60, No.8, pp.449-455 (2015).
- [18] 大澤翔吾, 村田剛志: 動的ネットワークにおける影響最大化, *人工知能学会論文誌*, Vol.30, No.6, pp.693-702 (2015).
- [19] 池田圭佑, 岡田佳之, 榊 剛史, 鳥海不二夫, 風間一洋, 野田五十樹, 篠田孝祐, 諏訪博彦, 栗原 聡: マルチエージェントによるマルチバースト型デマ拡散モデルの構築, *人工知能学会合同研究会 2014 データ指向構成マイニングとシミュレーション研究会 (DOCMAS)*, 一般社団法人人工知能学会 (2014).
- [20] 池田圭佑, 榊 剛史, 鳥海不二夫, 栗原 聡: 東日本大震災時のデマ情報拡散の分析, ネットワークが創発する知能研究会国内ワークショップ (JWEIN2016), 日本ソフトウェア科学会, 一般社団法人人工知能学会データ指向構成マイニングとシミュレーション研究会 (DOCMAS) 共催 (2016).
- [21] 池田圭佑, 榊 剛史, 鳥海不二夫, 風間一洋, 野田五十樹, 諏訪博彦, 篠田孝祐, 栗原 聡: マルチエージェント型情報拡散モデルの提案, *人工知能学会論文誌*, Vol.31, No.1, p.NFC-C.1 (2016).
- [22] 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).

付 録

A.1 パラメータの決定方法とパラメータ変更が結果に及ぼす影響

本付録では、AIDM で用いるパラメータを決定するために行った予備実験について述べる。なお、パラメータ決定時のデータが失われていたため、再度予備実験を行った。これ以降はそのデータを用いる。

A.1.1 予備実験の方法

今回実施した予備実験で、変更したパラメータは、「しきい値」、「忘却率」、「エージェントが状態遷移した際に MoT をどれだけ残すか (以降, MoT を残す割合)」の 3 つである。本予備実験ではこれら 3 つのうち 2 つのパラメータのみを固定し、それぞれ 500 回ずつシミュレーションを実施した。そして、各パラメータを決定する際はシミュレーション結果の平均情報発信数を参考とした。また、予備実

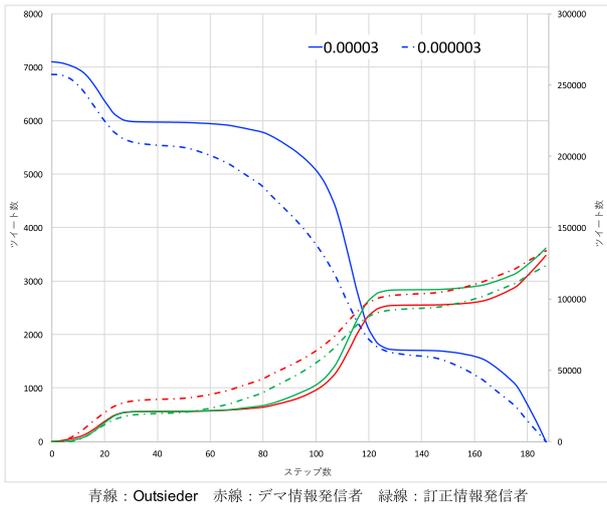


図 A.1 実験に用いたしきい値とその10分の1のしきい値の比較図 (左側縦軸: 0.00003, 右側縦軸: 0.000003)

Fig. A.1 Influence of threshold value on information diffusion: Left vertical axis 0.00003, Right vertical axis 0.000003.

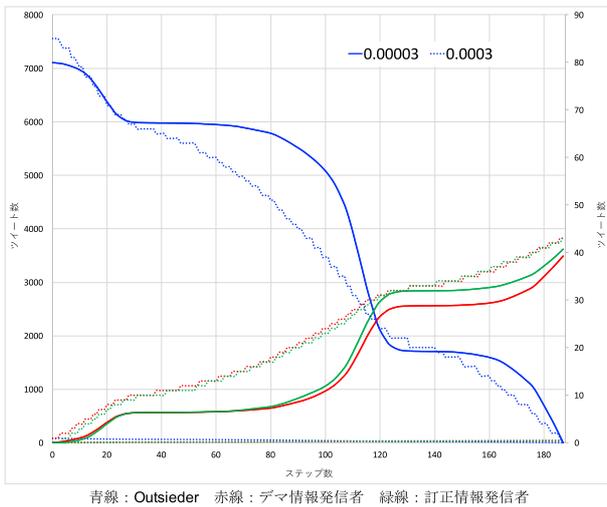


図 A.2 実験に用いたしきい値とその10倍のしきい値の比較図 (左側縦軸: 0.00003, 右側縦軸: 0.0003)

Fig. A.2 Influence of threshold value on information diffusion: Left vertical axis 0.00003, Right vertical axis 0.0003.

験では、コスモ石油に関するデマ情報の再現を行った。

それでは、各パラメータの変化が結果にどのような変化を及ぼすのかについて具体的に述べる。

A.1.2 しきい値を変更した場合 (忘却率および MoT を残す割合は論文と同様)

予備実験でのしきい値は、0.00003 (本稿で採用した値)、10分の1 (0.000003)、10倍 (0.0003) の3種類である。

なお、しきい値によって情報発信者数が大きく異なっており、結果を一度に図示できないため、本稿で採用したしきい値と別のしきい値という組合せで2種類の図を用意した (図 A.1, 図 A.2)。図 A.1, 図 A.2 より、しきい値によって大きく情報発信者数が異なることが分かる。

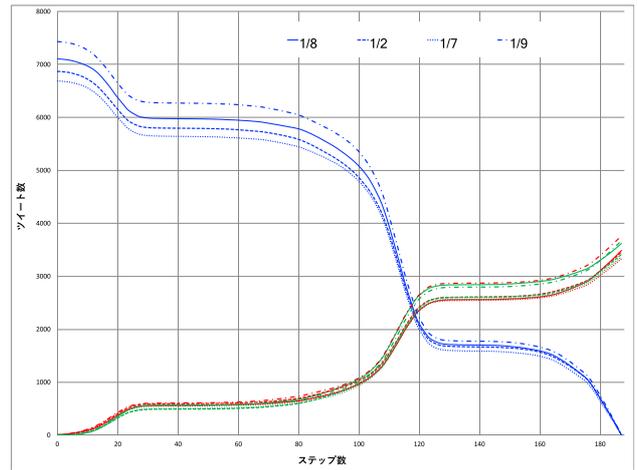


図 A.3 忘却率の変化による実験結果の変化

Fig. A.3 The influence of forgetting rate on information diffusion.

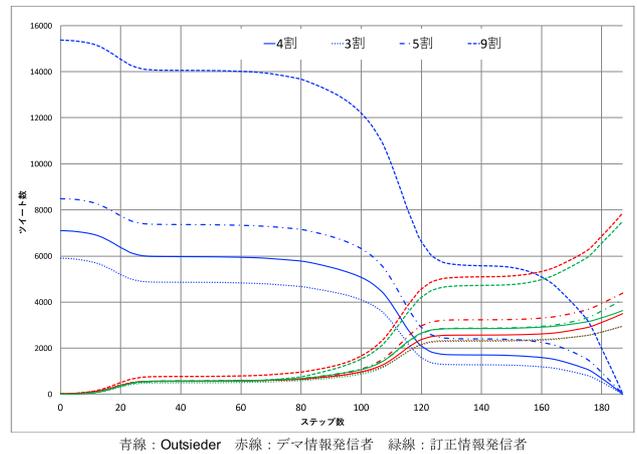


図 A.4 MoT を残す割合の変化による実験結果の変化

Fig. A.4 Influence of "MoT reset rate" change on information diffusion.

A.1.3 忘却率を変更した場合 (しきい値および MoT を残す割合は論文と同様)

予備実験での忘却率は、1/8 (本稿で採用した値)、1/2、1/7、1/9 の4種類である。

結果を図 A.3 に示す。図 A.3 より、忘却率の変化は、大きく実験結果に影響を及ぼさなかったことが分かった。

A.1.4 MoT を残す割合を変更した場合 (しきい値および忘却率は論文と同様)

予備実験での忘却率は、4割 (本稿で採用した値)、3割、5割、9割 の4種類である。

結果を図 A.4 に示す。図 A.4 より、MoT 大きく残すほどを情報発信者の総数が増えることが分かった。

A.1.5 予備実験のまとめ

本予備実験の結果より、情報発信者数の変化に大きく影

響を及ぼすのは、「しきい値」と「MoTを残す割合」であり、「忘却率」の変更は大きく情報発信者数に影響を及ぼさないことが分かる。



池田 圭佑

2013年電気通信大学電気通信学部電子工学科卒業。2015年同大学院情報システム学研究科社会知能情報学専攻博士前期課程修了。同年より同研究科博士後期課程に在学中。2016年より日本学術振興会特別研究員。ソーシャルメディア上での情報拡散現象の研究に従事。



榎 剛史

2004年東京大学工学部電子情報工学科卒業。2006年同大学院修士課程修了。電力会社通信部門での勤務を経て、2009年同大学院博士課程入学。2014年博士課程修了。博士（工学）。東京大学での特任研究員を経て、2015年より現職ならびに東京大学客員研究員。専門は、Webマイニング、自然言語処理、計算社会科学。



鳥海 不二夫（正会員）

2004年東京工業大学大学院理工学研究科機械制御システム工学専攻博士課程修了。同年名古屋大学情報科学研究科助手。2007年同助教。2012年東京大学大学院工学系研究科准教授。エージェントベースシミュレーション、ソーシャルメディア、計算社会科学、人工知能技術の社会応用などの研究に従事。人工知能学会、電子情報通信学会、日本社会情報学会各会員。博士（工学）。



栗原 聡（正会員）

慶應義塾大学大学院理工学研究科修了。NTT基礎研究所、大阪大学大学院情報科学研究科/産業科学研究所を経て、2012年より電気通信大学大学院情報理工学研究科教授。同大学人工知能先端研究センターセンター長。博士（工学）。ダウンゴ人工知能研究所客員研究員、ホンダR&D-Xアドバイザー、人工知能、複雑ネットワーク科学、ユビキタスコンピューティング等の研究に従事。著書『社会基盤としての情報通信』（共立出版）、『人工知能とは』（近代科学社）。翻訳『群知能とデータマイニング』、『スモールワールド』（東京電機大学出版）等。人工知能学会理事・編集長を歴任。人工知能学会、電子情報通信学会、日本ソフトウェア科学会、ACM各会員。