

マイクロブログにおける流言の特徴分析

宮部 真衣^{1,a)} 梅島 彩奈² 灘本 明代² 荒牧 英治^{1,3}

受付日 2012年4月20日, 採録日 2012年10月10日

概要: Facebook や Twitter など, マイクロブログの普及により, ユーザは様々な情報を瞬時に取得することができるようになった. 一方, マイクロブログでは流言も拡散されやすい. 流言は適切な情報共有を阻害し, 場合によっては深刻な問題を引き起こす. これまで, マイクロブログ上の流言拡散に関する分析は十分に行われていない. 本論文では, マイクロブログ上で拡散された流言の特徴について分析した. 分析対象は Twitter とし, 平常時の流言と災害時の流言を用いた. 分析の結果, まず, (1) マイクロブログにおける流言の特徴として, 以下の3点を明らかにした. (a) マイクロブログにおける流言の発信は, 不特定多数に向けたリツイートをインフラとしており, 口伝えでの流言の拡散と比較して, 流言内容の変容が起きにくい可能性がある. (b) マイクロブログ上では, 流言に関するツイートを複数回投稿するユーザは少なく, また流言情報を投稿したユーザが, 自身の発言を訂正することも少ない. (c) 流言に関する訂正ツイートが発信されても, 流言ツイートはすぐには収束しない. また, (2) マイクロブログにおける, 災害時と平常時の流言の違いとして, 平常時は連鎖的に流言が広がり, 災害時は1つのツイートが連鎖することなく爆発的に広がる傾向があることを明らかにした. これらは, 今後マイクロブログ上での流言拡散への対策を講じるうえでの重要な知見となる.

キーワード: マイクロブログ, 流言, 災害

Characteristic Analysis of Rumor and Correction Texts on Microblog

MAI MIYABE^{1,a)} AYANA UMEJIMA² AKIYO NADAMOTO² EIJI ARAMAKI^{1,3}

Received: April 20, 2012, Accepted: October 10, 2012

Abstract: Microblogging systems such as Twitter and Facebook have enabled the exchange of information in real time, allowing people to obtain not only useful information but potentially harmful information, such as rumors. Rumors block adequate information sharing and cause serious problems. Several studies have already analyzed rumors, but it remains unclear how rumors are spread on microblogging systems. This paper presents a case study of how rumors are spread on Twitter. We used tweets in both emergency and nonemergency situations. The results revealed four findings: (1) Rumors are mainly transmitted as re-tweets on Twitter. Therefore, rumors spread on Twitter may have fewer incidences of changes in content. (2) People seldom transmit multiple tweets about a rumor. Moreover, people who transmit a tweet about a rumor seldom transmit a correction tweet. (3) Many people continue to transmit tweets about rumors, even if someone transmits a correction tweet. (4) In a nonemergency situation, rumors tend to spread as a chain reaction. In an emergency situation, on the other hand, rumors tend to spread explosively without a chain reaction. Findings (1), (2), and (3) are trends characteristic of rumors on Twitter, and finding (4) shows the difference between an emergency and a nonemergency situation. We believe that these findings can provide important fundamental knowledge applicable toward countermeasures to help eliminate rumors on microblogging systems in the future.

Keywords: microblogging system, rumor, emergency situation

¹ 東京大学知の構造化センター
Center for Knowledge Structuring, The University of Tokyo,
Bunkyo, Tokyo 113-8656, Japan

² 甲南大学知能情報学部
Faculty of Intelligence and Informatics, Konan University,
Kobe, Hyogo 658-8501, Japan

³ 科学技術振興機構さきがけ
Japan Science and Technology Agency (JST) PRESTO,
Chiyoda, Tokyo 102-0076, Japan

^{a)} mai.miyabe@gmail.com

1. はじめに

ふだん人々は、情報伝達手段としてテレビ・電話・メールなどを頻繁に利用している。これらのメディアは大きく2つに分けることができるといわれている。テレビ・電話・チャットのような情報を発信するタイミングと閲覧するタイミングが同じものは「同期型メディア」、新聞・手紙・メールのようにタイミングが異なるものは「非同期型メディア」と呼ばれている。そして近年、これら双方の特徴を持った「選択同期型メディア」が登場した[1]。「選択同期型メディア」とは、ユーザが同期・非同期を自由に選択できるメディアのことであり、Facebook^{*1}やTwitter^{*2}などのマイクロブログは選択同期型メディアに該当する。新たなメディアの登場により、ユーザはより自由に情報交換を行うことが可能となった。

主に自身の状況や雑記などを短い文章で投稿するマイクロブログは、ユーザの情報発信への敷居が低く[2]、現在、マイクロブログを用いた情報発信が活発に行われている。2011年3月11日に発生した東日本大震災においては、緊急速報や救難物資要請など、リアルタイムに様々な情報を伝える重要な情報インフラの1つとして活用された[3], [4], [5]。マイクロブログは、重要な情報インフラとなっている一方で、情報漏洩や流言の拡散などの問題もかかえている。実際に、東日本大震災においても、様々な流言が拡散された[6]。本論文では、マイクロブログの問題の1つである、流言に着目する。流言は適切な情報共有を阻害する。特に災害時には、流言が救命のための機会を損失させたり、誤った行動をとらせたりするなど、深刻な問題を引き起こす場合もある。そのため、マイクロブログ上での流言の拡散への対策を検討していく必要があると考えられる。

マイクロブログの代表的なツールとして、Twitterがある。Twitterは、投稿する文章（以下、ツイート）が140字以内に制限されていることによる情報発信の敷居の低さと、リツイート（RT）という情報拡散機能により、流言が拡散されやすくなっている。実際に、東日本大震災においては、Twitterでは様々な流言が拡散されていたが、同じソーシャルメディアであっても、参加者全員が同じ情報と意識を持ちやすい構造を採用しているmixi^{*3}やFacebookでは深刻なデマの蔓延が確認されていないという指摘もある[7]。

マイクロブログ上での流言の拡散への対策を検討するためには、まずマイクロブログ上の流言の特徴を明らかにする必要がある。そこで本論文では、マイクロブログ上で拡散された流言の特徴について分析する。分析の対象として、

東日本大震災時にも多くの流言が拡散されていたTwitterを扱う。また、災害時のデータ（2011年3月、東日本大震災時）と平常時のデータ（2010年3月）を用いる。

本研究のポイントは以下の2点である。

- (1) 従来の流言に関する研究（現実社会の中での流言に関する研究）での知見やTwitterの特徴に基づき、マイクロブログ上での特徴を分析する。
- (2) 災害時のデータ（2011年3月、東日本大震災時）と平常時のデータ（2010年3月）を用いて、それぞれの違いの有無を検証する。

以下、2章において関連研究について述べる。3章では分析の概要について述べる。4章で分析結果を示し、5章で分析結果およびマイクロブログ上での流言に関する課題について考察する。最後に6章で本論文の結論についてまとめる。

2. 関連研究

本論文では、平常時および災害時のマイクロブログ上での流言について分析を行う。そこで本章では、まず、流言に関するこれまでの定義について述べた後、災害や流言について扱ったソーシャルメディアに関する研究について述べる。

2.1 流言の定義と流言の伝達

流言については、これまでに多くの研究が多方面からなされている。流言と関連した概念として噂、風評、デマなどの研究がある。これらの定義の違いについては諸説あり、文献ごとにゆれているのが実情である。本研究では、根拠がない情報を流言と定義し、その発生過程（捏造か自然発生か）は問わないものとする。

流言の分類としては、ナップによる第2次世界大戦時の流言の分類がある[8]。ナップは、流言を「恐怖流言（不安や恐れへの投影）」「願望流言（願望への投影）」「分裂流言（憎しみや反感への投影）」の3つに分類している。また、これらの流言がどの程度の割合で流通するかは社会状況によって決まると述べられている。社会状況は流言を伝達させる要因の1つであり、たとえば震災の直後など、社会状況が多くの人々に不安を感じさせる状況は、流言の発生や伝達に関係する。

また、流言の伝達には、曖昧さ、重要さ、不安という3つの要因が強く関係することが示されている[8]。オルポートとポストマンは、流言の流布量について、 $R \sim i \times a$ のように定式化し、「流言の流布量(R)は、重要さ(i)と曖昧さ(a)の積に比例する」と述べている[8]。

しかし、これらの流言に関する先行研究は、現実社会の中での口伝えでの流言の伝達について行われたものであり、マイクロブログでの流言の伝達に関して、十分な分析は行われていない。本論文では、マイクロブログを対象と

^{*1} <http://www.facebook.com/>

^{*2} <http://twitter.com/>

^{*3} <http://mixi.jp/>

し、先行研究での知見とマイクロブログの特徴に基づき、マイクロブログ上の流言の特徴について分析する。

2.2 災害、流言とソーシャルメディア

本節では、災害、流言を扱った、Twitter をはじめとするソーシャルメディアの研究について述べる。

Twitter をセンサとしてとらえ、災害などの異常事態の検出を試みた研究がある。Sakaki らは、Twitter を用いた地震や台風の位置の推定に関する研究を行っている [9]。Abel らは、緊急放送システムをモニタリングしておき、災害の発生を確認した後、Twitter から災害に関連するツイートを収集し、有益な情報をユーザに提供するシステムの開発を行っている [10]。Aramaki らや Paul らは、Twitter を用いてインフルエンザの把握を行っている [11], [12]。これらは、平常時からソーシャルメディアなどを監視しておくことで、異常事態発生時にいち早くその情報を伝えるという警告型のサービスである。

災害時のソーシャルメディアの利用について、定量的に分析した研究も多い。災害時にソーシャルメディア上に表出する感情の変化を扱った研究として、Back らや Cohn らによる研究がある [13], [14]。Back らや Cohn らは、9.11 時のブログの書き込み内容を分析し、人々の感情の変化を分析している。災害時のソーシャルメディアの利用方法について分析した研究としては、Longueville らや Qu ら、Vieweg らの研究がある [15], [16], [17], [18]。Longueville らは、2009 年にフランスで発生した森林火災に関して、Twitter に発信されたツイートの分析を行っている [15]。この研究においては、ツイートの発信者の分類や、ツイートで引用された URL の参照内容に関する分析などを行っている。Qu らは四川大地震および青海地震において中国のオンラインフォーラム (BBS) がどのように利用されたのかを分析している [16], [17]。Vieweg ら [18] は、2009 年のオクラホマの火事 (Oklahoma Grassfires) やレッドリバーでの洪水 (Red River Floods) における Twitter の利用方法を調査している。これらの研究では発信された内容を分類し、情報の発信の方法 (情報発信か返信か) や、その位置関係について議論しているが、情報が流言かどうかといった観点からの分析は行われていない。

また、近年では、流言の検知など、ソーシャルメディア上の流言を扱った様々な研究も行われている。Qazvinian らは、マイクロブログ (Twitter) における特定の流言に関する情報を網羅的に取得することを目的とし、流言に関連するツイートを識別する手法を提案している [19]。Mendoza らは、2010 年のチリ地震における Twitter ユーザの行動について分析を行っている [20]。この研究では、正しい情報と流言に関するツイートを、「支持」「否定」「疑問」「不明」に分類し、支持ツイート、否定ツイートの数について、正しい情報と流言との違いを分析している。分析結果として、

正しい情報を否定するツイートは少ないが (0.3%)、流言を否定するツイートは約 50% にのぼることを示している。

本論文では、災害時の流言とあわせて平常時の流言についても分析し、災害時と平常時との違いの有無を検証する。

3. 分析の概要

本章では、分析に用いるデータセットおよびツイートの定義について述べる。

3.1 対象データセット

本研究では、分析対象のデータとして、以下の 2 種類のデータを用いる。

平常時データ 2010 年 3 月^{*4}に投稿されたツイートのうち、流言キーワードを含むもの (一部の流言データについては、2010 年 2 月のツイートも含む)、7,232 件

災害時データ 2011 年 3 月 11 日~30 日^{*5}までに投稿されたツイートのうち、流言キーワードを含むもの、48,967 件

これらのデータの中から、以下の手順で流言を選定し、流言に関するツイートを抽出した。

- (1) データ全体から、情報の真偽について言及しているツイートをキーワード「デマ」をもとに抽出する^{*6}。
- (2) 手順 (1) で抽出したツイートから、無作為に 1,000 件抽出する。
- (3) 抽出した 1,000 件のツイートを確認し、間違いを含む情報 (流言) を分析候補として抽出する^{*7}。
- (4) 分析候補である流言を特定可能なキーワード (流言キーワード) をもとに、データ全体からツイートを抽出する。
- (5) 流言キーワードを含むツイートが 100 件以上^{*8}存在する流言を分析対象とする。

上記の手順により、平常時データについては 3 種類 (平常時 A~C)、災害時データについては 5 種類 (災害時 A~E) の流言に関するツイートを抽出し、分析対象とした^{*9}。

^{*4} 平常時のデータについては、暫定的に災害時データの 1 年前に設定した。

^{*5} 災害時のデータについては、災害発生直後に発生した流言を分析するために、2011 年 3 月 11 日をデータ抽出開始日とし、データ抽出終了日は暫定的に 3 月 30 日とした。データを抽出した結果、各流言がこの期間内にほぼ収束していることが確認できたため、災害時のデータについては、データ取得期間を 3 月 11 日~30 日とした。

^{*6} 今回は、真偽を言及する際に「デマ」というキーワードが用いられる可能性があると考え、「デマ」を含むツイートを抽出した。

^{*7} 本論文の流言の定義に基づき、「間違いを含む情報である」と十分に判断してよいことを、文献 [6]、ニュースの情報をもって確認した。

^{*8} 母数の少ないものを除外するため、暫定的に閾値を 100 件とした。

^{*9} 平常時データと災害時データの分析対象となる流言数が異なるのは、選定手順における選定条件を満たした分析対象の数が異なるためである。選定手順により得られた分析対象は、ランダムサンプリングによって得られる流言であると見なし、今回は分析対象数を調整せず、得られた結果をそのまま用いることとした。

表 1 データセット

Table 1 Examples of rumor tweets.

流言データ	流言内容	ツイート例
平常時	A アニメ D の声優交代	テレビ局 A は 25 日にホームページ上でアニメ「アニメ D」の声優を視聴率低迷のため変更することを発表した。O (73) さんをはじめとする旧レギュラー声優が復帰する予定。
	B テレビ局 N での Twitter 禁止	社内ツイッター禁止情報 テレビ局 N は一昨日より twitter.com への社内からのアクセスを禁止にした模様。
	C Twitter での歌詞のつぶやき	RT @****: Twitter で歌詞をつぶやくと組織 J の利用料が発生する。
災害時	A サーバルームでの負傷	地震が起きた時、社内サーバルームにいたのだが、ラックが倒壊した。腹部を潰され、血が流れている。
	B 命の三角形	命の三角形、ためになりました。地震では机の下にすぐ入る事を考えがちだけど、机等のすぐ横のほうがつぶされない三角形の空間になり、助かる可能性が高いんだね。
	C K 電力による節電呼びかけ	【関西地区の皆さん】K 電力が電力の提供を始めたようなので、コチラで節電すれば立派な支援になります。出来るかぎり節電を心がけましょう。あと、電子レンジや炊飯器等を使っていないときはコンセント等を抜いておくと電気節約&もし地震が来た際、火災防止にもなります。
	D O 氏の寄付	漫画 O の作者 O 氏、地震の被害者救済に 15 億円を寄付 「自分が幸せになったということは、世の中から受けたひとつの借りだ」
	E S 氏のラッキー発言	「K さんはとてもラッキーな人」M 党の S 前・官房長官が東北大地震を「ラッキー」と表現。

平常時の流言データについては、2010 年 3 月のツイートデータ (2010 年 2 月のデータも一部含む) のうち、各流言のキーワードを含むツイートを抽出し、作成した。災害時の流言データについては、2011 年 3 月 11 日 16 時 10 分から 3 月 30 日 17 時 20 分までに投稿された日本語ツイートから、各流言のキーワードを含むツイートを抽出し、作成した。なお、表中の流言内容およびツイート例は、実在の名称をイニシャルで提示している。

表 2 流言キーワード

Table 2 Keywords in rumor tweets.

流言データ	流言キーワード
平常時	A アニメ D, 声優
	B テレビ局 N, 社内, Twitter, アクセス, 禁止
	C Twitter, 歌詞, 組織 J, 利用料
災害時	A 社内サーバルーム
	B 命の三角形
	C K 電力, (節電, 給電, 送電, 供給, 提供, 依頼, 願い)
	D O 氏, (寄付, 15 億)
	E S 氏, ラッキー

表中の括弧で括られたキーワードは、いずれか 1 つを含むものを抽出した (たとえば、災害時 D の場合は、「O 氏 AND 寄付」または「O 氏 AND 15 億」を含むツイートを抽出した)。

表中のキーワードは、実在の名称をイニシャルで提示している。

分析に用いた流言の内容およびツイート例を表 1 に、流言キーワードを表 2 にそれぞれ示す。

3.2 ツイートの分類

本論文では、抽出した流言に関するツイートを以下の 3 種類に分類し、分析を行う。

(1) 流言ツイート：流言に関する発言

(2) 疑問ツイート：流言に対して、疑問を表す発言 (例：○○○は本当なの？ デマじゃないの?)

(3) 訂正ツイート：流言であることを指摘する発言 (例：このツイートはデマです。RT xxx: ○○○)

情報 (流言) の信頼性について疑問を表す表現を含むものは疑問ツイート、明らかに訂正であると分かるものは訂正ツイートとして扱う。

3 種類のツイートへの分類を人手で行った結果を表 3 に示す。表 3 より、訂正や疑問の発信数が最も多かった流言は災害時 D (訂正と疑問が全体の約 73%)、最も少なかった流言は平常時 B (訂正と疑問が全体の約 12%) であった。

上記の分類と同様の分類を行った研究として、Mendoza らの研究がある [20]。Mendoza らは、2010 年のチリ地震における Twitter 上の流言の分類を行い、流言を否定するツイート (本研究における訂正ツイート) が約 50% にのぼることを示している。しかし、文献内で提示された各流言について、訂正と疑問が占める割合を確認したところ、最も多い場合は 90%、最も少ない場合は 9% となっており、流言によって傾向は異なる。これは、上述した我々の分類結果と同じ傾向である。

なお、すべての分析対象データの分類は、1 名の判定者により行っている。上記の判定者を含む 6 名の判定者により 435 ツイートの分類作業を行い、 κ 係数を用いて一致率を算出した結果、 $\kappa = 0.876$ となり、十分に高い一致率が確認できた。判定者による判定結果は他の判定者ともおおむね一致すると判断し、今回は 1 名の判定者による判定結果を用いて分析を行う。

表 3 流言ツイート数, 疑問ツイート数, 訂正ツイート数

Table 3 Number of rumor tweets, question tweets, and correction tweets.

		流言 ツイート数	疑問 ツイート数	訂正 ツイート数	合計
平常時	A	589 (57.0%)	119 (11.5%)	326 (31.5%)	1,034 (100%)
	B	1,350 (87.9%)	60 (3.9%)	125 (8.1%)	1,535 (100%)
	C	3,934 (84.3%)	196 (4.2%)	533 (11.4%)	4,663 (100%)
災害時	A	2,822 (73.4%)	38 (1.0%)	984 (25.6%)	3,844 (100%)
	B	2,836 (87.4%)	34 (1.0%)	374 (11.5%)	3,244 (100%)
	C	15,883 (46.1%)	808 (2.3%)	17,791 (51.6%)	34,482 (100%)
	D	1,886 (27.1%)	169 (2.4%)	4,906 (70.5%)	6,961 (100%)
	E	267 (61.2%)	32 (7.3%)	137 (31.4%)	436 (100%)

表中の数値は各流言に関して抽出したツイートを3種類(流言, 疑問, 訂正)に分類した結果であり, それぞれの中にリツイート(非公式リツイートおよび公式リツイート)やリプライを含む。

3.3 訂正形式の分類

3.2 節で提示した訂正ツイートについて, 先行研究における定義をもとに, さらに分類を行った。

先行研究では, 噂への対処戦略として, 以下の3種類があると述べられている [8]。

否定戦略: 噂の内容を明確に否定する。

対抗戦略: 噂について否定しない。噂自体とは異なるイメージを流す。

無視戦略: 噂に対して反応せず, 噂が流れるままにしておく。

これらの対処戦略のうち, 無視戦略が行われたかどうかを判断することは難しいが, 否定戦略と対抗戦略が行われたかどうかは調査可能である。

ユーザは, 噂の内容を否定する際に, 「デマ」「間違い」「嘘」のような流言であることを示す表現(以下, 流言マーカと呼ぶ)を用いる可能性がある。流言マーカを含む訂正ツイートは, 否定戦略と見なすことができる。本論文では, Twitter の特徴であるリツイートをふまえ, この流言マーカに注目して, 訂正ツイートを以下の5種類((1)-(a)~(d), (2))に分類する。

(1) 否定戦略: 流言内容を否定しているツイート

(a) 「流言内容」「流言マーカ」の順で出現するツイート

(例: **** (流言内容) は デマ です)

(b) 「流言マーカ」「流言内容」の順で出現するツイート

(例: デマ だったんだね, **** (流言内容))

(c) 「流言マーカ」「RT」「流言内容」の順で出現するツイート

(例: デマ だったようです RT **** (流言内容))

(d) 流言マーカを含まないが, 流言を否定しているツイート

(例: **** (流言内容) について, ネットではソースが見つからない)

(2) 対抗戦略: 正確な情報を含むツイート

なお, 本論文では, 17 語句^{*10}を流言マーカとし, 流言マーカの有無を確認し, ツイートを分類する。訂正ツイートの分類結果を表 4 に示す。

表 4 より, 各流言によって, 投稿される訂正ツイートの傾向が異なることが分かる。さらに, 対抗戦略についてはまったく発信されない場合(災害時 A, 災害時 D, 災害時 E)があるが, 否定戦略については, 発信数に違いはあるものの, いずれの流言においても発信されていることが分かる。

4. 分析結果

本章では, 表 3 に示した各ツイートをもとに, マイクロブログ上での流言の拡散および流言への対処について分析する。

流言拡散の対策を講じるためには, まず, 「流言がどのように拡散されているのか?」という点を明らかにする必要がある。従来研究では, 口伝えのコミュニケーションにおける流言の拡散傾向について分析されているが [21], 口伝えとマイクロブログでは, 流言拡散の性質が異なる可能性がある。そのため, マイクロブログ上の流言の拡散傾向を明らかにし, 従来の知見との違いについて分析する必要がある。また, もし, マイクロブログ上でユーザが自身の投稿した流言への対処を行っている場合, 新たに対策を講じる必要性がない可能性もある。従来研究では, 口伝えのコミュニケーションを対象として, 噂への対処戦略の定義や, 各対処戦略の効果に関する実験的な検証が行われているが [8], 流言を伝達した個人による流言への対処に関する分析は十分に行われていない。しかし, マイクロブログでは削除や訂正をしなければ, 投稿した流言はテキストとして残るため, 口伝えの場合と異なり, ユーザによる対処が重要となる可能性がある。そのため, マイクロブログ上での流言拡散に対し, ユーザ自身による十分な対応が行われているかどうかを調査する必要がある。また, 災害流言が深刻な影響を与えうることから, 従来研究においては災害流言が特に注目されてきたものの, 災害流言をもって流言全体を考えることは危険である, という指摘が

^{*10} 本論文で用いた流言マーカは, 「デマ」「嘘」「ウソ」「釣り」「ツリ」「偽情報」「都市伝説」「ガセ」「ネタ」「狂言」「迷信」「誤報」「間違い」「いたずら」「騙され」「釣られ」「チェーンメール」である。

表 4 訂正形式の分類
Table 4 Classification of correction tweets.

		否定戦略				対抗戦略	合計
		(a)	(b)	(c)	(d)		
平常時	A	171 (52.5%)	15 (4.6%)	117 (35.9%)	12 (3.7%)	11 (3.4%)	326 (100%)
	B	23 (18.4%)	2 (1.6%)	47 (37.6%)	4 (3.2%)	49 (39.2%)	125 (100%)
	C	0 (0%)	1 (0.2%)	3 (0.6%)	0 (0%)	525 (99.2%)	529 (100%)
災害時	A	53 (5.4%)	72 (7.3%)	289 (29.4%)	570 (57.9%)	0 (0%)	984 (100%)
	B	48 (12.8%)	5 (1.3%)	9 (2.4%)	280 (74.9%)	32 (8.6%)	374 (100%)
	C	5,247 (29.5%)	1,814 (10.2%)	106 (0.6%)	2,868 (16.1%)	7,756 (43.6%)	17,791 (100%)
	D	2,852 (58.1%)	1,362 (27.8%)	330 (6.7%)	362 (7.4%)	0 (0%)	4,906 (100%)
	E	55 (40.1%)	59 (43.1%)	14 (10.2%)	9 (6.6%)	0 (0%)	137 (100%)

ある [21]. つまり, 現実社会において拡散される流言については, 災害流言と平常時の流言には違いがある可能性があると考えられ, マイクロブログ上の分析についても, この点に考慮すべきであると考えられる.

そこで本論文では, 以下の 3 つのリサーチクエスチョン (RQ) を立てて検証する.

RQ1: 流言はどのような形で発信されるのか?

RQ2: 流言を拡散したユーザは, 自身の投稿に対処する
のか?

RQ3: 平常時と災害時では, どのような違いがあるか?

これらを通して, マイクロブログ上での流言拡散傾向を明らかにすることは, 従来の口伝えでの流言拡散との違いに関する新たな知見を与え, またマイクロブログ上の流言対策の基礎的な知見となると考えられる.

以降の節において, それぞれのリサーチクエスチョンについて検証する.

4.1 RQ1: 流言はどのような形で発信されるのか?

本節では, マイクロブログ上で発信された流言の傾向 (RQ1) について検証する.

まず, Twitter の機能であるリツイートとリプライに注目する. Twitter には, 「リツイート*¹¹」と呼ばれる他者のツイートの引用機能や, 「リプライ」と呼ばれる他者のツイートに対する返信機能がある. そこで, 流言ツイート, 疑問ツイート, 訂正ツイートに関して, それぞれリツイ

ートとリプライがどの程度含まれているかを調査した. 調査の際, ツイート内容に「RT @ユーザ名」が含まれるツイートおよび他のツイートと内容が完全一致したツイートをリツイート, ツイート内容が「@ユーザ名」で始まるツイートをリプライと見なした.

リツイート数およびリプライ数を表 5 に示す. 表 5 より, リツイート率は最大 98.9% (災害時 A の流言ツイート), 最小 14.5% (災害時 C の疑問ツイート) である. 特に, 流言ツイートについては, 最大 98.9% (災害時 A), 最小 74.2% (災害時 E) となっている.

一方, リプライ率は最大 13.7% (災害時 C の疑問ツイート), 最小 0% であり, 平常時 B, 災害時 C の疑問ツイートを除き, すべて 7% 未満にとどまっている. すなわち, 多くの流言ツイートは, 特定のユーザに向けて発信される (リプライ) のではなく, 不特定多数のユーザが閲覧可能な形 (リツイート) で発信されていることが分かる.

次に, 発信に影響しうる特徴的な表現に注目する. 「拡散希望」のような拡散を促す表現 (以下, 拡散フレーズと呼ぶ) は, 流言の発信に影響する可能性がある. そこで, 拡散フレーズを含むツイート数について調査した. 拡散フレーズ*¹²を含むツイート数を表 6 に示す.

表 6 より, 拡散フレーズ出現率は最大 27.0% (災害時 D の訂正ツイート), 最小は 0% である. ツイート全体の拡散フレーズ出現率を見ると, 平常時 A, 平常時 B, 平常時 C, 災害時 B については, いずれも 1% 未満にとどまっており, 拡散フレーズがない場合でも, 多くのユーザによって流言が発信されている. 災害時に着目すると, 拡散フレーズの

*¹¹ リツイートには, Twitter 運営側が公式機能として提供している公式リツイートと, リツイート機能が提供されるまでのローカルルールが適用された非公式リツイートがある. 本研究では非公式リツイートおよび公式リツイートをあわせてリツイートとして扱う.

*¹² 本論文では, 「拡散」「RT 希望」「RT してください」「要 RT」を拡散フレーズと見なし, 調査を行った.

表 5 リツイート数およびリプライ数
Table 5 Number of re-tweets and replies.

		リツイート数				リプライ数			
		流言	疑問	訂正	合計	流言	疑問	訂正	合計
平常時	A	466 (79.1%)	68 (57.1%)	212 (65.0%)	746 (72.1%)	8 (1.4%)	3 (2.5%)	12 (3.7%)	23 (2.2%)
	B	1,200 (88.9%)	43 (71.7%)	106 (84.8%)	1,349 (87.9%)	26 (1.9%)	5 (8.3%)	2 (1.6%)	33 (2.1%)
	C	2,973 (75.6%)	149 (76.0%)	276 (51.8%)	3,398 (72.9%)	58 (1.5%)	3 (1.5%)	9 (1.7%)	70 (1.5%)
災害時	A	2,791 (98.9%)	31 (81.6%)	951 (96.6%)	3,773 (98.2%)	46 (1.6%)	0 (0%)	20 (2.0%)	66 (1.7%)
	B	2,574 (90.8%)	17 (50.0%)	296 (79.1%)	2,887 (89.0%)	21 (0.7%)	1 (2.9%)	8 (2.1%)	30 (0.9%)
	C	12,315 (77.5%)	117 (14.5%)	13,864 (77.9%)	26,296 (76.3%)	771 (4.9%)	111 (13.7%)	1,031 (5.8%)	1,913 (5.5%)
	D	1,633 (86.6%)	91 (53.8%)	4,261 (86.9%)	5,985 (86.0%)	45 (2.4%)	8 (4.7%)	125 (2.5%)	178 (2.6%)
	E	198 (74.2%)	21 (65.6%)	105 (76.6%)	324 (74.3%)	12 (4.5%)	2 (6.3%)	7 (5.1%)	21 (4.8%)

表中に括弧で示した割合は、提示したリツイート数またはリプライ数を、表 3 に示した値で割ったものである。

表 6 拡散フレーズ出現数
Table 6 Number of tweets with “spreading” phrases.

		流言	疑問	訂正	合計
平常時	A	0 (0%)	0 (0%)	4 (1.2%)	4 (0.4%)
	B	1 (0.1%)	0 (0%)	0 (0%)	1 (0.1%)
	C	1 (0%)	0 (0%)	0 (0%)	1 (0%)
災害時	A	502 (17.8%)	8 (21.1%)	84 (8.5%)	594 (15.5%)
	B	22 (0.8%)	1 (2.9%)	0 (0%)	23 (0.7%)
	C	3,533 (22.2%)	23 (2.8%)	1,330 (7.5%)	4,886 (14.2%)
	D	14 (0.7%)	1 (0.6%)	1,324 (27.0%)	1,339 (19.2%)
	E	0 (0%)	1 (3.1%)	21 (15.3%)	22 (5.0%)

表中に括弧で示した割合は、提示した拡散フレーズ出現数を、表 3 に示した値で割ったものである。

出現率が高い流言と低い流言のどちらも存在し、大量に拡散される流言が拡散フレーズをとまうとは限らないことが分かる。

以上の結果から、マイクロブログ上で発信された流言の傾向をまとめる。マイクロブログにおける流言の発信は、不特定多数に向けたリツイートをインフラとしている。また、災害時においては、拡散フレーズの有無にかかわらず、流言は発信・拡散される。

4.2 RQ2：流言を拡散したユーザは、自身の投稿に対処するのか？

本節では、流言を拡散してしまったユーザによる、自身の投稿への対処の有無（RQ2）について分析する。

流言を拡散したユーザが自身の投稿への対処を行っている場合は、その流言を話題としたツイートを複数回投稿していると考えられる。そこで、まず、それぞれの流言を話題として、複数回ツイートを投稿するユーザがいるかどうかを確認した。各流言データについてツイートを投稿したユニークユーザ数を表 7 に示す。

表 7 より、複数回ツイートを投稿したユーザは、最大 15.4%（災害時 C）、最小 3.3%（災害時 A）であった。つまり、多くのユーザは、ある流言を含むツイートを 2 回以上投稿しないと考えられる。

しかし、少数ではあるが、ある流言に関するツイートを複数回投稿しているユーザも存在する。では、このようなユーザのうち、流言を投稿したユーザは、自身のツイートを訂正しているのだろうか？

そこで、複数回投稿したユーザに着目し、流言に対する状態がどのように変化したかを考察する。複数回投稿したユーザの流言に関する最古のツイート（初期状態）と、最新のツイート（最終状態）を確認し、流言を投稿してからその訂正を試みたかどうかを調査する。

複数回投稿ユーザの流言に対する状態の変化を表 8 に示す。なお、表中では流言ツイートを **R**、疑問ツイートを **Q**、訂正ツイートを **C** と表記することとし、ユーザの初期状態と最終状態のペアを **R-C**（初期状態：流言、最終状態：訂正）のように表現する。本論文では、3 種類のツイー

表 7 ユニークユーザ数
Table 7 Number of unique users.

流言データ		ユニークユーザ数				複数回投稿したユーザ数
		全体	流言ツイート	疑問ツイート	訂正ツイート	
平常時	A	885	566	114	301	124 (14.0%)
	B	1,348	1,243	61	101	129 (9.6%)
	C	4,241	3,676	196	494	309 (7.3%)
災害時	A	3,702	2,739	37	956	124 (3.3%)
	B	2,932	2,754	33	327	261 (8.9%)
	C	28,325	14,495	764	14,905	4,363 (15.4%)
	D	6,385	1,847	167	4,597	468 (7.3%)
	E	374	232	31	116	36 (9.6%)

表 8 複数回投稿ユーザの流言に対する状態の変化
Table 8 Change of reaction about rumors of users.

		R-R	R-Q	R-C	Q-R	Q-Q	Q-C	C-R	C-Q	C-C
平常時	A	11 (8.9%)	3 (2.4%)	75 (60.5%)	2 (1.6%)	3 (2.4%)	17 (13.7%)	0 (0%)	0 (0%)	13 (10.5%)
	B	78 (60.5%)	10 (7.8%)	33 (25.6%)	1 (0.8%)	0 (0%)	5 (3.9%)	0 (0%)	1 (0.8%)	1 (0.8%)
	C	181 (58.6%)	5 (1.6%)	92 (29.8%)	14 (4.5%)	0 (0%)	8 (2.6%)	0 (0%)	0 (0.8%)	9 (2.9%)
災害時	A	69 (55.6%)	1 (0.8%)	26 (21.0%)	1 (0.8%)	1 (0.8%)	1 (0.8%)	0 (0%)	1 (0.8%)	24 (19.4%)
	B	64 (24.5%)	8 (3.1%)	155 (59.4%)	0 (0%)	0 (0%)	3 (1.1%)	5 (1.9%)	5 (1.9%)	21 (8.0%)
	C	871 (20.0%)	47 (1.1%)	1,295 (29.7%)	30 (0.7%)	20 (0.5%)	142 (3.3%)	152 (3.5%)	41 (0.9%)	1,765 (40.5%)
	D	27 (5.8%)	2 (0.4%)	179 (38.2%)	3 (0.6%)	2 (0.4%)	29 (6.2%)	2 (0.4%)	4 (0.9%)	220 (47.0%)
	E	18 (50.0%)	0 (0%)	2 (5.6%)	0 (0%)	1 (2.8%)	2 (5.6%)	0 (0%)	0 (0%)	13 (36.1%)

表中の R は流言ツイート, Q は疑問ツイート, C は訂正ツイートを意味する. 括弧中の割合は, 各流言の複数回投稿したユーザ全体に占める割合である.

トを投稿した段階のユーザの状態は, 以下のように考えることとする.

流言ツイートを投稿した場合: ある情報(流言)が事実であると考えている状態

疑問ツイートを投稿した場合: ある情報(流言)が事実であるか疑問に思っている状態

訂正ツイートを投稿した場合: ある情報(流言)が流言であることを認識している状態

以下において, 複数回投稿ユーザに関する, 各状態からの状態変化について述べる. なお, 災害時 E については, 他の流言と比較して複数回投稿ユーザ数が少ないため, 参考値として提示する.

(1) 初期状態が R (流言ツイートを投稿)

表 8 より, 災害時 C~E を除き, 複数回投稿ユーザの 70% 以上は, 初期状態が R である. また, 初期状態が R である複数回投稿ユーザの割合が最も低いのは災害

時 D で 44.4% であり, 災害時 C, E については 50% 以上は初期状態が R である. つまり, 流言に関するツイートを複数回投稿しているユーザは, 初期状態として流言が事実であると考えたユーザが多いと考えられる. また, 最終状態を確認すると, 災害時 E を除いた流言については, いずれも 20% 以上のユーザが最終的には訂正ツイートを投稿している. 一方, 平常時 A, 災害時 D を除いた流言については, 20% 以上のユーザは流言ツイートを再度投稿しており, 状態変化が見られなかった.

(2) 初期状態が Q (疑問ツイートを投稿)

特徴的な現象として, Q-R というペアが, 平常時データに最大 4.5% 発生していた. つまり, 疑問ツイートを投稿後に, 流言ツイートを投稿するという現象が発生する場合もあることが分かる.

(3) 初期状態が C (訂正ツイートを投稿)

初期状態が C の場合の多くについては、最終状態も C であり、状態変化が見られない。また、特徴的な現象として、災害時 B~D 以外の流言については、C-R というペアが発生していない。C-R というペアが発生した流言について確認を行ったところ、災害時 B, D については、最終的なツイートの内容が、流言を話題にしているが、訂正や疑問と判定できない内容（災害時 B におけるツイート例：「命の三角形」の話は、中学校の時、防災訓練で教えてもらった記憶がある。そういう考えもあるという話と共に、しかし机にもぐること怪我を防げる場合もあるからと、通常の訓練を行った。いい先生だったな）であるため、C-R と判定されていた。また、災害時 C については、上述した災害時 B, D と同様の原因により C-R と判定されたもの以外に、電力会社による電力提供に関して、「提供するらしい」という情報や「提供が始まった」という情報が投稿されており、前者の情報を訂正するツイートを投稿した後、後者の情報を信じて流言ツイートを投稿したため、C-R と判定された場合があった。しかし、災害時 B~D についても C-R というペアの発生率は低い。したがって、訂正ツイート投稿後に、流言ツイートを投稿するという現象は発生しにくいことが分かる。

今回の分析対象データにおいては、流言ツイートを投稿したユーザの多くについて、状態変化は観測できなかった。ただし、今回は流言に関するキーワードを用いて流言に関するツイートを抽出している。そのため、たとえば、「さきほどのツイートは間違っていました」のような、流言に関するキーワードを含まない形で訂正ツイートを投稿している場合については検証できていない。この点を考慮すると、本論文で明らかにしたことは、「流言投稿者による、投稿した流言への対処は、流言の内容を明示した形（流言を特定可能なキーワードを含む形）ではほとんど行われていない」ということである。

以上の結果から、流言を拡散してしまったユーザによる、自身の投稿への対処の有無についてまとめる。流言情報を投稿したユーザが、流言の内容を明示したうえで自身の発言を訂正することは少ない。これは、マイクロブログ上において、同じ流言に関するツイートを複数回投稿するユーザは少なく、また、複数回投稿するユーザについても、最終的に訂正ツイートを投稿するユーザは多いとはいえないためである。

4.3 RQ3: 平常時と災害時では、どのような違いがあるか?

本節では、平常時と災害時の違い (RQ3) について分析する。

まず、3.3 節、4.1 節、4.2 節で示した内容について、平

表 9 引用関係に基づくツイートの定義

Table 9 Definition of “tweet” based on citation relation.

	他のツイートを引用していない	他のツイートを引用している
他のツイート内で引用されていない	単独ツイート	終端ツイート
他のツイート内で引用されている	ルートツイート	中継ツイート

常時と災害時の違いがあるかどうかを考察する。リツイート数やリプライ数、訂正の形式については、平常時と災害時との明確な違いは見られなかった。各流言によって傾向が異なるため、投稿された流言の内容などが影響している可能性がある。

一方、拡散フレーズの有無については、平常時と災害時で違いが見られた。4.1 節で示したように、平常時データのツイート中に拡散フレーズはほとんど含まれない。災害時データについては、拡散フレーズをほとんど含まない場合と、20%程度含む場合があった。拡散フレーズを含む割合の多いデータを見ると、災害時 A の内容は救助の要請、災害時 C の内容は節電の要請である。緊急性があり、人間の行動を促すような内容が含まれており、拡散を促進するために拡散フレーズが用いられた可能性がある。災害時には、このような内容を含むものが発生しやすく、拡散フレーズを含むツイートが存在したと考えられる。

次に、発信された流言や訂正の広がり方に違いがあるのかを検証する。4.1 節において、流言に関するツイートの多くは、リツイートとして発信されていることを示している。リツイートでは、何らかのツイートを引用している。発信されたリツイートをさらにリツイートすることも可能であり、リツイートが繰り返されることにより、流言が連鎖的に拡散していく可能性もある。

そこで、流言が広がっていく傾向を分析するために、本論文では、「他のツイートを引用しているか」および「他のツイートで引用されているか」という観点から、表 9 に示す 4 種類のツイート（ルートツイート、中継ツイート、終端ツイート、単独ツイート）に分類する。また、表 9 の定義に基づく引用関係の分類例を図 1 に示す。本論文では、リツイートを引用と見なすこととし、「RT @ユーザ名」を含むツイート（非公式リツイート）および他のツイートと内容が完全一致するツイート（公式リツイート）を引用関係の分析対象とする。

引用関係の抽出において、非公式リツイートについては、ツイート中に最初に出現する「RT @ユーザ名」に含まれるユーザのツイートを引用していると見なした。非公式リツイートにおける引用関係の抽出方法を以下に示す。

(1) 非公式リツイート中に最初に出現する「RT @ユーザ名」に含まれるユーザ名を被引用ユーザとして取得

する。

- (2) その非公式リツイートより前に投稿された、被引用ユーザのツイートのうち、非公式リツイート中で引用されている内容を含む最新のツイートを「引用されたツイート」とする。
- (3) 手順(2)において引用されたツイートが発見できなかった場合、今回用いたデータセット中に引用ツイートデータが存在しなかったと判断する(ただし、他のツイートの引用は行っているため、最終的に分類する際、その非公式リツイートがさらにリツイートされている場合は中継ツイート、そうでない場合は終端ツイートに分類する)。

公式リツイートについては、内容が完全一致するものは、すべて同じルートツイートを引用しているとした。非公式リツイート、公式リツイートの引用関係の判定例を図2に示す。リツイート(非公式リツイート、公式リツイート)は、中継または終端ツイートに分類される。公式リツイートは基本的には終端ツイートに分類されるが(図2のC)、公式リツイート(図2のD)が非公式リツイートされた場合(図2のF)、その公式リツイートは中継ツイートに分類される。リツイート以外のツイート(他の

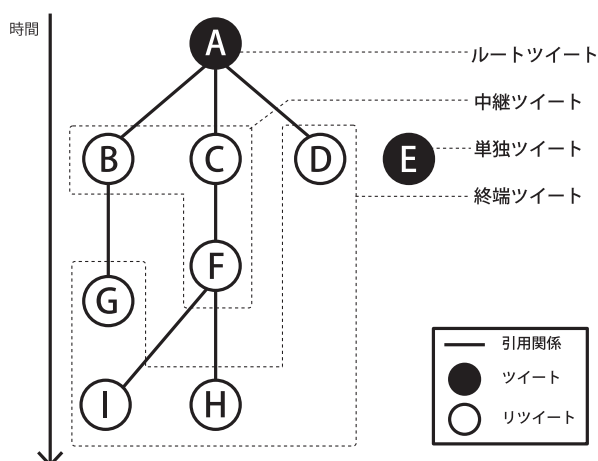


図1 引用関係の例

Fig. 1 Example of citation relation.

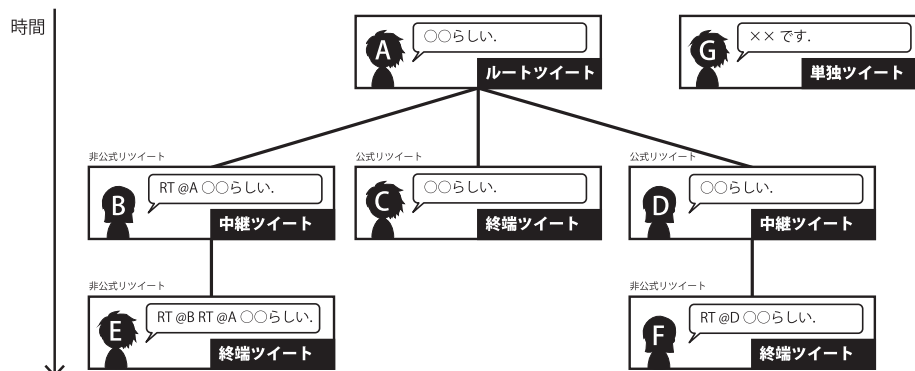


図2 非公式リツイート、公式リツイートの引用関係の例

Fig. 2 Example of judgment of citation relation.

ツイートと内容が完全に一致することがなく、「RT @ユーザ名」を含まないものはルートまたは単独ツイートに分類される。リツイート以外のツイートのうち、投稿後、他のツイートによって引用されたものはルートツイートに分類され、それ以外のは単独ツイートに分類される。

表9の定義に基づき、流言ツイートおよび訂正ツイートを分類した。それぞれの分類結果を図3および図4に示す。

図3より、流言の拡散においては、中継ツイートと終端ツイートの割合に関して、平常時データと災害時データと

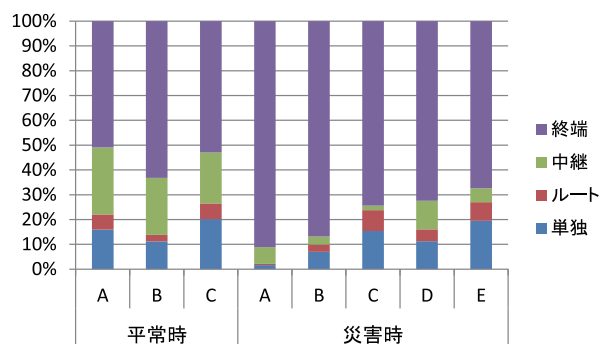


図3 流言ツイートの引用関係に基づく分類結果

Fig. 3 Classification result of rumor tweets based on citation relation.

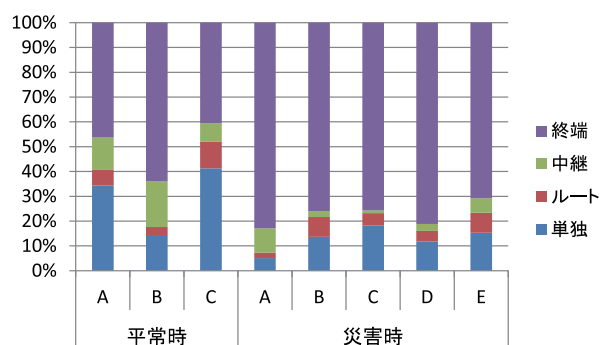


図4 訂正ツイートの引用関係に基づく分類結果

Fig. 4 Classification result of correction tweets based on citation relation.

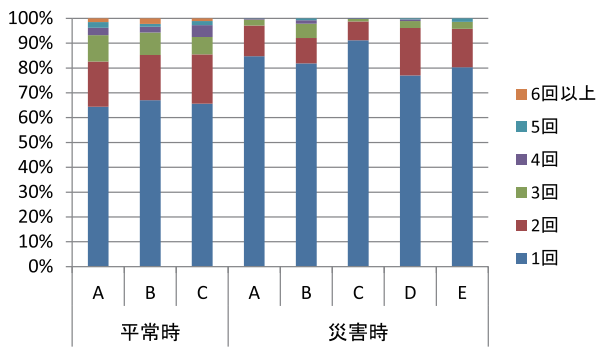


図 5 ルートツイートからの最大連鎖数

Fig. 5 Maximum number of chains of re-tweets.

の違いが見られる。まず、終端ツイートを見ると、平常時は最大 63.2%，最小 50.9%であるのに対し、災害時は最大 91.1%，最小 67.4%となっており、平常時よりも災害時の方が全体に占める割合が多い傾向が見られる。また、中継ツイートを見ると、平常時は最大 27.0%，最小 20.5%であるのに対し、災害時は最大 11.6%，最小 1.8%となっており、平常時よりも災害時の方が全体に占める割合が少ない。図 4 を見ると、平常時 B および災害時 A～E については、訂正の拡散においても、中継ツイートと終端ツイートに関して流言の拡散と同様の傾向が見られる。平常時 A と平常時 C については、流言の拡散と比較して、中継ツイートの減少と単独ツイートの増加が見られる。表 4 を見ると、平常時 A は否定戦略 (a) が、平常時 C は対抗戦略が全体に占める割合が多い傾向が見られる。3.3 節で述べたように、否定戦略 (a) は流言情報を短く要約しているツイートが多い。また、対抗戦略は正確な情報を含むツイートである。つまり、否定戦略 (a) や対抗戦略については、流言内容をそのまま引用せず、ユーザ自身が新たに入力している場合が多いと考えられる。そのため、単独ツイートの割合が増加し、中継ツイートの割合が減少した可能性がある。

中継ツイートは、リツイートが連鎖している（あるツイートを引用したリツイートが、さらにリツイートされている）ことを意味する。つまり、災害時は、中継ツイートが少なく、終端ツイートが多いことから、連鎖的に広がるのではなく、何らかのルートツイートが爆発的に広がる傾向があると考えられる。一方、平常時は災害時よりも中継ツイートが多いことから、連鎖的に流言が広がる傾向があると考えられる。

また、ルートツイートからの最大連鎖数を調査した。図 5 に、ルートツイートから末端ツイートまでの最大連鎖数の分布を示す。図 5 より、災害時の流言は、いずれも最大連鎖数 1 回が約 80%を占めている。また、6 回以上の連鎖は見られない。一方、平常時の流言は、いずれも最大連鎖数 3 回以上の流言の割合が災害時と比較して多い。また、いずれも 6 回以上の連鎖が発生しており、平常時 B については、最大 11 回の連鎖が発生していた。

以上の結果から、平常時と災害時の違いについてまとめる。平常時と災害時では、流言の広がり方に違いが見られる。具体的には、平常時は連鎖的に流言が広がり、災害時は 1 つのツイートが連鎖することなく爆発的に広がる傾向がある。

5. 考察

5.1 現実社会における流言拡散との違い

本節では、4 章で分析した内容をもとに、現実社会における流言拡散との違いについて述べる。

現実社会の中での流言は、基本的には 1 人の人間が、1 人あるいは複数人の人間へと伝達するというプロセスを繰り返すことにより、拡散されていく。従来研究で対象とされているものは、主として口伝えのコミュニケーションであり、伝達過程において、平準化、強調、同化といった現象が生じ、流言内容が変容すると述べられている [21]。

一方、マイクロブログ上の情報伝達はテキストにより行われる。Twitter では、リツイート機能により、テキストを引用することができる。リツイートでは、引用内容に対するコメントが付加されたり、文字数制限に基づき一部が削除されたりする場合もあるが、引用内容は基本的には改変されず、そのまま引用される。つまり、マイクロブログ上で拡散される流言は、リツイートをインフラとしていることから (4.1 節)、従来研究で示された、現実社会の中での流言の拡散と比較して、流言内容の変容が起こりにくい可能性がある。

5.2 マイクロブログ上の流言対処に関する課題

本節では、分析結果から明らかとなったマイクロブログ上の流言対処に関する課題について述べる。

マイクロブログ上での流言がどのように収束していったのかを確認するために、各ツイート数の時系列変化を調査した。図 6 に、流言ツイートおよび訂正ツイートの時系列変化のグラフを示す^{*13}。各グラフの横軸は、日付および時刻であり、災害時 A 以外は、1 時間単位である（災害時 A のみ、短時間で収束しているため、1 分単位とした）。また、各グラフの左側の縦軸は流言ツイートの件数、右側の縦軸は訂正ツイートの件数である。

図 6 より、平常時 A、平常時 B、災害時 B～E については、訂正ツイートが発生するタイミングは流言によって異なるものの、流言ツイートの増減と訂正ツイートの増減の流れが類似している。平常時 C については、流言ツイートが増加した際に、訂正ツイートがあまり増加せず、流言ツイートの増加が何度か発生し、徐々に減少していく様子が見られる。災害時 A については、流言ツイートが増加した際に訂正ツイートも増加する。その後、流言ツイートはす

^{*13} 疑問ツイートについては、各流言に占める割合が小さく、特徴的な変化が見られないため、グラフ中には提示していない。

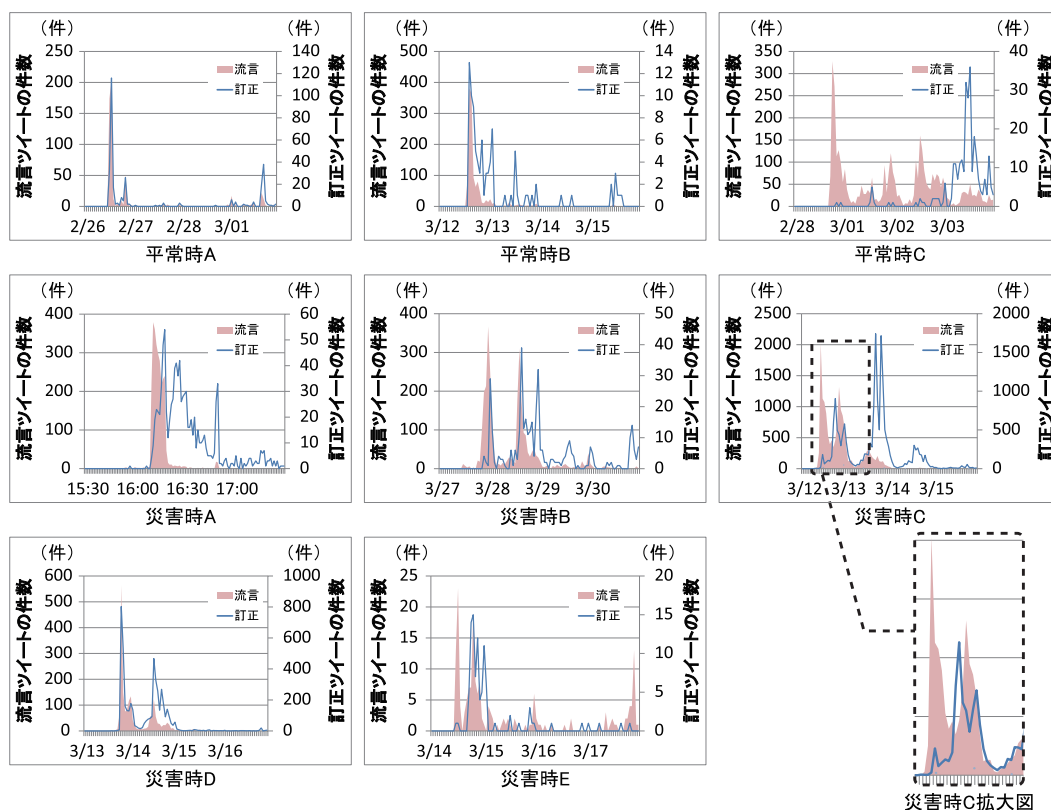


図 6 流言に関するツイートの時系列変化

Fig. 6 Time-line chart of rumor tweets.

ぐに収束し、訂正ツイートはしばらく投稿されるという現象が見られる。ただし、収束までの時間が短い（1 時間で流言ツイート、訂正ツイートが急速に増加し、収束している）ため、1 分ごとの推移を提示しており、1 時間単位で見た場合、平常時 A、平常時 B、災害時 B～E と同様に、流言ツイートの増減と訂正ツイートの増減の流れが類似していると考えられる。

一般に、人々がある情報を他者に伝える場合、その情報が正しいと思って伝えていることが多く、本人がでたらめだと思ふ話を、悪意をもって他者に伝えることは少ない [8]。つまり、流言の伝達は、主に伝達している情報が流言であることを認識していないことに起因すると考えられる。訂正ツイートがユーザの目に入れば、ユーザはその情報が流言である可能性に気づく可能性がある。ユーザが流言である可能性に気づけば、流言の拡散を防ぐことも可能となる。

しかし、いずれの流言についても、流言ツイートが投稿され始めた段階で訂正ツイートが投稿されているが、訂正ツイートの出現以降にすぐに収束するわけではなく、流言ツイートの投稿は続いていることが分かる。たとえば、図 6 の災害時 C 拡大図を見ると、訂正ツイートが出現した後も、流言ツイートの投稿が続いている。4.2 節で述べたように、流言を投稿したユーザの多くは、自身の投稿内容に対処しない。そのため、そのユーザのツイートを閲覧している他のユーザには、訂正ツイートが行きわたるとは

限らず、流言の拡散が続いたと考えられる。

はじめに述べたように、流言は適切な情報共有を阻害し、場合によっては深刻な問題を引き起こす。そこで、本論文で明らかにした特徴をもとに、流言の拡散を防ぐための方針についてまとめる。マイクロブログ上での流言拡散における問題点としては、流言を拡散したユーザが自身の発言を訂正することは少なく（4.2 節）、訂正ツイートが発信されても流言ツイートはすぐには収束しない（本節）という点がある。この問題については、訂正ツイートが発信された場合に、その流言ツイートの投稿者および閲覧者に対し、流言である可能性を提示するといった対策が可能であると考えられる。今後、このような対策を実現するための仕組みについて検討していく必要がある。

5.3 本研究の適用限界

流言の分析結果は、分析するデータの影響を受ける可能性がある。本節では、本論文における分析の限定性についてまとめる。

本研究では、分析データの期間を限定し、流言の抽出を行った。取得するデータ期間によって、分析対象とした Twitter ユーザ数や、利用のされ方などが異なる可能性がある。また、分析対象とした流言は、選定手順によって抽出した 8 種類であり、本研究で分析した流言とは異なる傾向を示す流言も存在すると考えられる。

本研究で抽出した分析対象に対しては、本論文の流言の定義に基づき、「間違いを含む情報である」と十分に判断してよいことを、文献やニュースなどの情報^{*14}をもって確認した。しかし、対象期間のツイートから2章で定義した流言をすべて抽出できたわけではない。また、分析対象とする流言を選定する手順においては、「デマ」というキーワードを用いており、今回分析対象とした流言は、3.3節で述べた否定戦略を少なくとも1件は含むデータである。そのため、対抗戦略のみが発生し、否定戦略がまったく発信されていない流言は、今回分析対象としたデータには含まれていない。

また、4.2節で述べたように、今回分析した訂正ツイートには、「さきほどのツイートは間違っていました」のような、流言キーワードを含まない形の訂正ツイートは含まれない。このような形での訂正ツイートは、流言ツイートの投稿から時間をあけずに投稿されている場合、どの流言に対する訂正なのかを適切に関連付けできる可能性がある。しかし、時間が経過してから訂正を投稿されている場合、間に他のツイートを投稿されていたり、他のユーザのツイートが投稿されたりすることにより、訂正がどの流言に対するものなのかを正しく関連付けできなくなる可能性がある。そのため、他のユーザに対して正しく訂正情報を伝達するためには、流言の内容が明示されていることが重要であると考え、今回は、流言キーワードを含む訂正ツイートについて分析を行った。しかし、流言キーワードを含まない訂正ツイートが、流言の拡散・収束に影響を及ぼしている可能性もある。

今後は、異なる分析データの期間における流言を抽出することにより、多様な流言の分析し、データ取得期間による影響について検証する必要がある。また、上述した選定手順における制約を外し、今回の分析対象に含まれない流言や訂正を抽出、分析することにより、マイクロブログ上で発生しうる流言の傾向をより明確にできる可能性がある。

6. おわりに

本研究では、マイクロブログ上での流言拡散の特徴を明らかにするために、Twitterを例とした分析を行った。分析対象として、平常時のデータ（2010年2月、3月）およ

び災害時のデータ（2011年3月11日～30日）から抽出した、8種類の流言を用いた。流言に関するツイートを3種類（流言ツイート、疑問ツイート、訂正ツイート）に分類し、マイクロブログ上での流言に対する反応について確認した。分析の結果、マイクロブログ上の流言拡散には、以下の傾向があることを明らかにした。

(1) マイクロブログにおける流言の特徴

- マイクロブログにおける流言の発信は、不特定多数に向けたリツイートをインフラとしており、口伝えでの流言の拡散と比較して、流言内容の変容が起こりにくい可能性がある（5.1節）。
- マイクロブログ上では、流言に関するツイートを複数回投稿するユーザは少なく、また流言情報を投稿したユーザが、自身の発言を訂正することも少ない（4.2節）。
- 流言に関する訂正ツイートが発信されても、流言ツイートはすぐには収束しない（5.2節）。

(2) マイクロブログにおける、災害時と平常時の流言の違い

- 平常時は連鎖的に流言が広がり、災害時は1つのツイートが連鎖することなく爆発的に広がる傾向がある（4.3節）。

今後は、これらの知見に基づき、マイクロブログ上での流言拡散を防ぐための仕組みを検討していく必要がある。

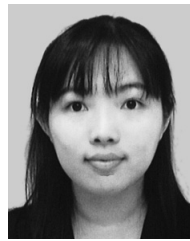
謝辞 本研究のデータ分析において、貴重かつ膨大なデータ（平常時）を提供していただいた兼山元太氏（クックパッド）に深く感謝する。本研究の一部は、JST 戦略的創造研究推進事業による。

参考文献

- [1] 濱野智史：アーキテクチャの生態系—情報環境はいかに設計されてきたか、エヌティティ出版（2008）。
- [2] 垂水浩幸：実世界インタフェースの新たな展開：4. ソーシャルメディアと実世界、情報処理学会誌，Vol.51, No.7, pp.782-788（2010）。
- [3] インプレス R&D インターネットメディア総合研究所：インターネット白書 2011，インプレスジャパン（2011）。
- [4] 西谷智広：I 見聞録：Twitter 研究会，情報処理学会誌，Vol.51, No.6, pp.719-724（2010）。
- [5] 立入勝義：検証 東日本大震災 そのときソーシャルメディアは何を伝えたか？，ディスカヴァー・トゥエンティワン（2011）。
- [6] 荻上チキ：検証東日本大震災の流言・デマ，光文社新書（2011）。
- [7] 小林啓倫：災害とソーシャルメディア—混乱，そして再生へと導く人々の「つながり」，毎日コミュニケーションズ（2011）。
- [8] 川上善郎：うわさが走る—情報伝搬の社会心理，サイエンス社（1997）。
- [9] Sakaki, T., Okazaki, M. and Matsuo, Y.: Earthquake shakes Twitter users: Real-time event detection by social sensors, *Proc. 19th International Conference on World Wide Web (WWW '10)*, pp.851-860（2010）。
- [10] Abel, F., Hauff, C., Houben, G.J., et al.: Twitcident:

^{*14} 平常時 A については、2010年3月の時点で、ツイートの記載された内容が発生していないことから、間違いであると判断した。平常時 B については、テレビ局 N の宣伝部の公式アカウントによる訂正情報の投稿が行われていたことから、間違いを含む情報であると判断した。平常時 C についてはニュース記事 (<http://www.itmedia.co.jp/news/articles/1003/19/news017.html>) をもとに、間違いであると判断した。災害時 A, C, D, E については、文献 [6] により間違いであることを確認した。災害時 B については、命の三角形に関する情報 (<http://www.snopes.com/inboxer/household/triangle.asp>) に対し、反論記事 (<http://www.cert-la.com/RejoinderToDougCopp.pdf>) が存在しており、必ずしも正しい情報ではないことから、間違いを含む情報であると判断した。

- Fighting Fire with Information from Social Web Stream, *Proc. International Conference on Hypertext and Social Media*, pp.305-308 (2012).
- [11] Aramaki, E., Maskawa, S. and Morita, M.: Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter, *Proc. 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP2011)*, pp.1568-1576 (2011).
 - [12] Paul, M.J. and Dredze, M.: You Are What You Tweet: Analyzing Twitter for Public Health, *Proc. 5th International AAAI Conference on Weblogs and Social Media (ICWSM 2011)*, pp.265-272 (2011).
 - [13] Back, M.D., Kufner, A.C.P. and Egloff, B.: The Emotional Timeline of September 11, 2001, *Psychological Science*, Vol.21, No.10, pp.1417-1419 (2010).
 - [14] Cohn, M.A., Mehl, M.R. and Pennebaker, J.W.: Linguistic markers of psychological change surrounding September 11, 2001, *Psychological Science*, Vol.15, No.10, pp.687-693 (2004).
 - [15] Longueville, B.D., Smith, R.S. and Luraschi, G.: "OMG, from here, I can see the flames!": A use case of mining location based social networks to acquire spatio-temporal data on forest fires, *Proc. 2009 International Workshop on Location Based Social Networks (LBSN '09)*, pp.73-80 (2009).
 - [16] Qu, Y., Wu, P.F. and Wang, X.: Online Community Response to Major Disaster: A Study of Tianya Forum in the 2008 Sichuan Earthquake, *42nd Hawaii International International Conference on Systems Science (HICSS-42 2009)*, pp.1-11 (2009).
 - [17] Qu, Y., Huang, C., Zhang, P., et al.: Microblogging after a major disaster in China: A case study of the 2010 Yushu earthquake, *Proc. ACM 2011 Conference on Computer Supported Cooperative Work (CSCW '11)*, pp.25-34 (2011).
 - [18] Vieweg, S., Hughes, A.L., Starbird, K., et al.: Microblogging during two natural hazards events: What twitter may contribute to situational awareness, *Proc. 28th International Conference on Human Factors in Computing Systems (CHI '10)*, pp.1079-1088 (2010).
 - [19] Qazvinian, V., Rosengren, E., Radev, D.R., et al.: Rumor has it: Identifying Misinformation in Microblogs, *Proc. 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP2011)*, pp.1589-1599 (2011).
 - [20] Mendoza, M., Poblete, B. and Castillo, C.: Twitter under crisis: Can we trust what we RT?, *Proc. 1st Workshop on Social Media Analytics (SOMA '10)*, pp.71-79 (2010).
 - [21] 早川洋行: 流言の社会学—形式社会学からの接近, 青弓社 (2002).



宮部 真衣 (正会員)

1984 年生。2006 年和歌山大学システム工学部デザイン情報学科中退。2008 年同大学大学院システム工学研究科システム工学専攻博士前期課程修了。2011 年同大学院システム工学研究科システム工学専攻博士後期課程修了。博士 (工学)。現在、東京大学知の構造化センター特任研究員。コミュニケーション支援に関する研究に従事。



梅島 彩奈

1989 年生。2012 年甲南大学知能情報学部知能情報学科卒業。在学中は、Web コミュニケーション技術の習得や、マイクロブログ上の流言の特徴分析に関する研究に従事。



灘本 明代 (正会員)

会員

東京理科大学理工学部電気工学科卒業。2002 年神戸大学大学院自然科学研究科後期博士課程修了。博士 (工学)。現在、甲南大学知能情報学部教授。Web コンピューティング、データ工学の研究に従事。ACM, IEEE 各



荒牧 英治 (正会員)

1974 年生。2000 年京都大学総合人間学部卒業。2002 年同大学大学院情報学研究科修士課程修了。2005 年東京大学大学院情報理工系研究科博士課程修了。博士 (情報理工学)。現在、東京大学知の構造化センター特任講師、科学技術振興機構さきがけ研究員 (兼任)。自然言語処理応用の研究に従事。