

災害時 Twitter における デマとデマ訂正 RT の傾向

梅島 彩奈[†] 宮部 真衣^{††}
荒牧 英治^{††} 灘本 明代[†]

Facebook や Twitter 等、マイクロブログの普及により、ユーザはインターネット上でリアルタイムなコミュニケーションを容易に行う事が可能になっている。特に Twitter は 2008 年に国内でのサービスを開始して以降、急速にユーザ数を伸ばしている。災害時において、電話・手紙・メールを使用できない場合でも、Twitter は使用することができる。このように Twitter は多くの人々と様々な情報を共有できる半面、デマの拡散を助長することもある。そこで本論文では、ツイートのデマとデマの訂正に注目し、これらの拡散に関する 6 つの仮説を立て、その仮説に基づきデマとデマ訂正ツイートの傾向を分析することで、デマの拡散を防止することを目指す。

Tendency of Rumor and Correction Re-tweet on the Twitter during Disasters

AYANA UMEJIMA[†] MAI MIYABE^{††}
EIJI ARAMAKI^{††} AKIYO NADAMOTO[†]

The rapid progress of the micro blog such as Facebook and Twitter, user can communicate with each other easily in real time on the Internet. Especially, fast growing Twitter has so many users. We could use twitter during the disaster in Japan, even if we could not use telephone, cell-phone, and e-mail. In this way, the twitter becomes popular information tools. There, however, much rumor information on the twitter and user can spread it easily. In this paper, we focus on the re-tweet of rumor and correction. First, we propose 5 hypothesis, next we analysis the tweets during the disaster in Japan to substantiate our proposed hypothesis.

1. はじめに 【*の文字書式「隠し文字」】

Facebook や Twitter 等、マイクロブログの普及により、利用者は各々が別の場所に居ながらにして、リアルタイムなコミュニケーションを可能としている。特に 140 字以内のツイートと呼ばれる文章を投稿し、閲覧できるコミュニケーションサービスである Twitter は、2008 年に国内でのサービスを開始して以降、2011 年 6 月現在 1 ヶ月間の国内利用者数が 1,290 万人を超え、1 日の国内ツイート数は 8 万ツイートを突破している。

一方、普段人々はコミュニケーションツールとして、電話・手紙・メールを使用することが多いが、実際に 2011 年 3 月 11 日に起こった東日本大震災においてこれらのツールはほとんど役に立たず、Twitter は使用することが出来た。このように災害時のコミュニケーションツールとして Twitter は重要な位置づけとなってきている。

その反面、Twitter にはいくつかの問題がある。その一つとして挙げられるのが、情報の正確性である。Twitter には、リツイート (RT) という情報拡散機能がある。これにより、ユーザが必要だと判断した情報をより多くの人に届けることが出来る。しかし、RT するかどうかの判断はユーザに委ねられているため、時として間違った情報が拡散されてしまうこともある。例えば、「放射線対策にイソジン・ワカメが良い」の様なものである。このように間違った情報を本論文ではデマと呼ぶ。

非常時にデマと気づかずにデマ情報を取得することは、更なる混乱を招くことにつながる。ほとんどのユーザがツイートによるデマ発生当初はそのツイートが間違っていることには気づかずに善意で拡散してしまう。しかし、そのツイートが多くユーザに拡散されると、知識のあるユーザによってそのツイートは訂正される。これで、一見問題は解決したように見えるが、よく観察してみると、デマ RT はデマ訂正 RT に比べて遥かに多く拡散され続けている。これにより、正確な情報は膨大な量のデマによって隠されてしまう。このことは、災害時における情報収集において解決すべき最重要課題の一つといえる。

そこで我々は、ツイートのデマ情報を自動で取得できたら便利であると考え、本研究では、災害時の Twitter におけるデマとデマ訂正の拡散の傾向を分析し、ツイートのデマ情報を自動取得する手法の提案を行う。本論文ではそのはじめの一歩として、災害時の Twitter におけるツイートのデマとデマ訂正の分析を行う。分析を行うに当たり、まず始めに我々は本研究において、“デマは拡散しやすいが、デマの訂正は拡散しにくい”という仮説を立てた。

[†]甲南大学知能情報学部知能情報学科
Konan University
^{††}東京大学知の構造化センター
The University of Tokyo

本稿ではこの仮説を証明するために、まず実際のツイート进行分类し、そして分類結果の「情報発信系」のツイートに注目し、さらに5つの仮説を立てて、分析を行いRTされやすいツイートの特徴を抽出する。

以下、2章では関連研究について、3章ではツイートの分類について述べ、4章では、分類結果の「情報発信系」のツイートに見られる、いくつかの特徴をユーザ実験から抽出し、考察について述べ、5章でまとめと今後の課題について述べる。

2. 関連研究

2.1 デマについての関連研究

デマについては多くの研究が多方面からなされている。また、デマと関連した概念として噂／風評／流言などの研究がある。これらの定義の違いについては諸説あり、文献毎にゆれているのが実情である。たとえば、デマは悪意ある情報を捏造する現象とし、流言は自然発生的なものとした分類もある。また、真実という十分な根拠がない情報をデマとする、誤った情報をデマとするなどの違いもある。本研究では、デマとは、根拠がない情報をデマと定義し、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。(1)不安や恐れへの投影としての流言（恐怖流言）、(2)願望への投影としての流言（願望流言）、(3)憎しみや反感への投影としての流言（分裂流言）がある。

また、オルポート等は流言の発生を支えているのは、不安、不満下での緊張を緩和や他人がまだ知らない情報を他人に知らせたいという心理的メカニズムによるものとし、流言の伝播量は当事者に対する問題の重要性と、その問題についての根拠のあいまいさとの積に比例していると述べている。ロスノウらは「あいまいさ」のほかに流言の伝播に影響する要因として「不安」（その流言を聞いて不安を喚起される程度）や「信用度」（その流言の真実への確信度）をあげている。

本研究では、これらの分析をソーシャルメディアを用いて定量的に行う。

2.2 ソーシャルメディアと災害について

Twitterをはじめとしたソーシャルメディアと災害との関係については、近年、多くの研究がなされている。Sakaki等 [1]は、地震や台風の位置をTwitterを用いて推定した。Aramaki等 [2], Paul等 [3]はTwitterを用いてインフルエンザの把握を行っている。

これらの研究は通常時から、つねにソーシャルメディアをウォッチしておき、災害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という点ではアイデアを同じくする。

一方、災害時におけるソーシャルメディアの利用を定量的に分析した研究も多い。Back [4]やCohn等 [5]は、9.11時のブログの書き込みを分析し、人々の感情の変化を分析した。

Qu等 [6]は四川地震における中国のオンラインフォーラム（BBS）の利用のされ方を分析した。Vieweg等 [7]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）でのTwitterの利用を調査した。

これらの研究では情報の発信の方法（情報発信か返信か）や、その位置関係について議論されているが、情報がデマかどうかといった観点からの分析は行われていない。

3. TwitterRTの分類

3.1 TwitterRT

RTとは、Retweetの略で、他のユーザの投稿を再投稿することである。発言元のツイートがそのまま表示される「公式RT」と、ユーザが「RT @username 引用文」といった書式に編集する「非公式RT」がある。「公式RT」の場合、発言元ツイートの編集は行えない仕様となっている [8]。RTされることで、ツイートが多くユーザの目に触れる機会が増加する。

3.2 RTの分類

我々は、2011年3月11日から30日の間に収集されたツイートの内、RT数が多いもの上位1000件のRTを収集し、その分析を行った。

その結果、RTの内容は、「実用的な特徴か」「私情を含む情報か」「実話か」「経験情報を含むか」の4つの特徴をもとにして、5つに分類することができた（図1参照）。5つの分類の定義・ツイート例を以下に示し、分析結果の種類別件数と割合を表1に示す。

- **情報発信系**：災害時において実用的な情報であり、且つ、私情を含まず且つ経験情報を含まない。そして、ニュースや注意喚起等、大勢のユーザに伝達することを旨とした情報。
例)【先ほどの地震】おそらく全く同時刻で長野北部と関東で揺れを検出したため、大きな地震と誤報が出たようです。
- **経験談**：災害時において実用的な情報であり、且つ、私情を含まず且つ経験情報を含む。そして、災害被害に関する経験談や伝聞情報、それらを踏まえた助言。
例) 阪神地震の人からです！新聞紙とラップがかなり役に立つと。お腹に新聞紙を巻いて上からラップを巻く！熱を逃がさないそうです！ラップは他にも多様性が

あるのでかなり重宝するそうです！

- **私見**：災害時において実用的な情報であり、且つ、私情を含む。そして、ある情報について個人の意見が述べられているもの。
例) 地震以降、有名無名関係なく Twitter ユーザの誰もが 1 回は失言したと思うんですよ。判断ミスもあった。それを恥じる必要は無いっていうか、うん、恥ずかしいんだけど、みんなそんなもんなんですよ。
- **小話**：災害時において実用的な情報ではなく、且つ実話であるもの。災害被害以外の経験談・伝聞情報。
例) いまラジオで KARA が今回の地震に対して募金したとのニュースが流れたけど、給料5万なのに！
- **ジョーク**：災害時において実用的な情報ではなく、且つ実話ではないもの。ユーザを笑わせる目的で作られたもの、冗談。
例) 地震「ぼぼぼーん！」 俺「こっちくん」 地震「遊ぼう」 俺「お前とは友達になれないな」 地震「ばか」 俺「ばか」 地震「もう遊ばない」 俺「もう遊ばない」 そうやってあとでさみしくなって 地震「ごめんね」 俺「許さないよ」 地震「えっ台本と違う」 俺「お前が言うな」

これら5種類のツイートの中でも特に「情報発信系」のツイートは、1000件中749件と多くの割合を占める。さらに、不特定多数に向けられた情報や注意喚起、緊急性の高い情報を含み、ユーザに与える影響がある可能性が非常に高い。その為、これらのツイートの中にデマが含まれることは、ユーザを不安に陥れるだけでなく、時にはユーザに間違った行動を促す可能性が考えられる。

以上のことから、我々は本研究において「情報発信系」のツイートを対象とする。

4. 情報発信系 RT のデマに関する分析

4.1 情報発信系 RT の分類

我々は更に3章にて分類したツイートの内、情報発信系 RT を詳しく分析した。その結果、情報発信系 RT には、いくつかの特徴が見られることがわかった。

特徴とその詳細を以下に示し、分析結果の種類別件数と割合を表2に示す。

- **拡散**：他のユーザに RT を希望する旨を記載することで、ツイートの拡散を助長するもの。ここでは、「拡散希望」「RT してください」等と記載されていることが多い。

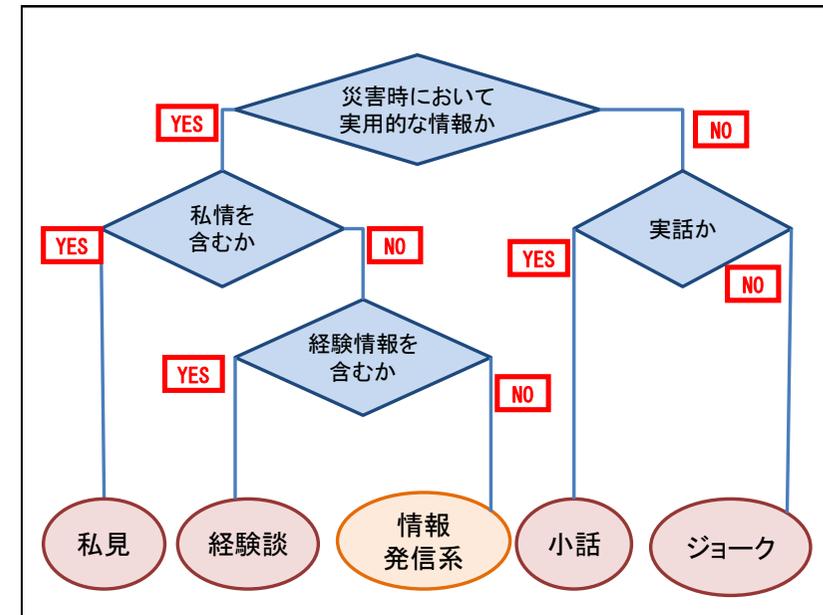


図 1 TwitterRT の分類フローチャート
Fig.1 Flow Chart of Classification of TwitterRT

表 1 TwitterRT の種類別件数
Table 1. Number of TwitterRT by Categories

種類	ツイート件数	割合(%)	RT 数	割合(%)
情報発信系	749	74.9	130711	62.32
経験談	36	3.6	8916	4.25
私見	82	8.2	22902	10.92
小話	100	10	39514	18.84
ジョーク	33	3.3	7706	3.67
合計	1000		209749	

- **URL**：そのツイートに関連する情報が得られる web ページの URL を記載しているもの。
- **ハッシュタグ**：そのツイートに関連するハッシュタグを記載しているもの。ハッシュタグとは、ツイートでキーワードや話題を明示的に表現するために用いられる#で始まる文字列のことである。#に続く文字列は半角英数字であり、ハッシュタグの前後にはスペースを必要とする。
- **詳細情報**：そのツイートに関する詳細情報を記載しているもの。詳細情報とは、情報の詳細を知るために必要な情報のこと。住所や電話番号等。

表 2 TwitterRT の特徴別件数
Table 2. Number of TwitterRT by Characteristics

特徴	全 RT 内件数	情報拡散系内件数	割合(%)
拡散	103	89	88.3
URL	361	333	92.2
ハッシュタグ	193	171	89.1
詳細情報	46	41	91.3

4.2 デマの特徴発見の為の評価属性

デマ特有の特徴を発見するために、ツイートの印象について分析した。その評価属性の定義とツイート例を以下に示す。

- **行動促進強**：そのツイートに影響され、すぐに行動しなければならないと感じる内容である。
例) 地震が起こったら、必ず窓を開けてください。そして、家にいる人は、今、お風呂に水をためてください。
- **行動促進弱**：すぐに行動するほどではないが、心を動かされる内容である。
例)【緊急地震速報について】3月11日の大震災の影響で地震計が被害を受けたため緊急地震速報の精度がかなり落ちています。大元は気象庁より発表されたものですが、震度で大幅な誤差が生じることが多いです。自動Postを鵜呑みにせず、テレビ等で正確な情報収集に努めてください。
- **ポジティブ**：前向きになる、喜ばしい内容である。

例) 今日。地震。津波。原発。で被災された方々にお見舞い申し上げます。復興に携わっている方々、自衛隊の方々、有難うございます。がんばれ日本。がんばれ日本。

- **ネガティブ**：後ろ向きになる、落ち込む内容である。
例)【生産活動】充電式乾電池を作っている国内最大の群馬県高崎市の工場では、地震後、電池の需要が高まっているにもかかわらず、計画停電の影響で生産量が2割ほど落ち込む状態が続いています。
- **不安を煽る**：不安になる内容である。
例) 阪神淡路大震災は三時間後に最大の揺れがきたらしいので、気を緩めないでください。

4.3 デマのツイートの仮説と実験

我々は、情報発信系 RT における4つの特徴と5つの評価属性をもとに、5つの仮説を立てた。

- 仮説 1. URL を含む RT はデマである可能性が低い
- 仮説 2. デマは「行動を促す」「ネガティブな」「不安を煽る」内容が多い
- 仮説 3. 「行動を促す」内容であるツイートは RT されやすい
- 仮説 4. 「ネガティブな」内容であるツイートは RT されやすい
- 仮説 5. 「不安を煽る」内容であるツイートは RT されやすい

そして、仮説を証明するために以下の実験を行った。

被験者は6名で、全員大学生である。実験には3.2節で分類した749件の情報発信系 RT を用いた。被験者には各々のツイートを見てもらい、4.2節の評価属性について各々評価してもらった。実験結果を示すとともに、以下で5つの仮説の証明を行う。

4.3.1 仮説 1. URL を含む RT はデマである可能性が低い

URL を含むツイートは情報源がはっきりしている。その為、デマである可能性が低いと考えられる。そこで、情報発信系 RT に含まれるデマと URL を含むツイートの件数を調べた(表 3 参照)。

なお、各々のツイートがデマであるかどうかの判定には、以下の WEB サイトの情報を参考にした。

- TwiDema[11]
- 松永英明 公式サイト 絵文録ことのは「震災後のデマ 80 件を分類整理して見えてきたパニック時の社会心理」[12]

- 荻上式 BLOG 「東北地方太平洋沖地震，ネット上でのデマまとめ」 [13]
- ついのすみか 「東北地方太平洋沖地震のデマ情報まとめ」 [14]

表 3 デマと URL 付きツイートの件数比較

Table 3. Comparison between Dema tweet and tweet with URL

	全体 (件)	URL 付き (件)	割合(%)
情報発信系	749	333	44.46
非デマ	715	323	45.17
デマ	30	10	33.33
デマ訂正	4	0	0

実験の結果，情報発信系の 44.46%が URL 付きであり，そのうち非デマは 45.17%，デマは 33.33%となり，デマにおける URL 付きツイート件数の割合は，デマでないもの（非デマ）に比べて低いことがわかる．つまりは，「URL を含む RT はデマである可能性が低い」事がわかった．しかし，デマにおける URL 付きツイート 10 件について詳しく見ていくと，URL 先の情報が個人の WEB ページであることがわかった．個人のページであるため，情報が必ずしも正確とは言えず，デマにつながったと考えられる．

4.3.2 仮説 2. デマは「行動を促す」「ネガティブな」「不安を煽る」内容が多い

本研究では，“流言の伝播に影響する要因である「不安」を「ネガティブ」「行動を促す」「不安を煽る」の 3 つの尺度に分けて，実験を行う．表 4 は，被験者が選択したツイートの RT 数の合計値とその割合を示している．

実験の結果，行動促進については，行動促進強がデマでない場合 34.33%であるのに対し，デマの場合は 77.31%と，顕著に比率が高いことがわかる．また，行動促進弱に関しても，デマでない場合が 82.96%に対し，デマの場合は 96.26%と高いことがわかる．さらに，「ポジティブ」の RT の割合は，デマでない場合 70.12%であるのに対し，デマの場合 66.97%と減少しているのに対し，「ネガティブ」はデマでない場合 51.30%に対しデマの場合は 91.06%となり顕著に高い事がわかる．また「不安を煽る」についてもデマでない場合は 48.04%であるのに対しデマの場合は 78.46%と高くなっている．このことより，「行動促進強・弱」「ネガティブ」「不安を煽る」の全てにおいて，デマは他に比べて割合が顕著に高いことが判明し，仮説 2 は正しいと言える．

表 4 RT 数と評価属性の比較

Table4. Comparison between Number of RT and attribute of evaluation

	行動促進強		行動促進弱	
	RT 数	割合(%)	RT 数	割合(%)
情報発信系 130711RT	48547	37.14	109581	83.84
非デマ 122456RT	42045	34.33	101587	82.96
デマ 6981RT	5397	77.31	6720	96.26

	ポジティブ		ネガティブ		不安を煽る	
	RT 数	割合(%)	RT 数	割合(%)	RT 数	割合(%)
情報発信系 130711RT	91810	70.24	70334	53.81	65573	50.17
非デマ 122456RT	85861	70.12	62815	51.30	58822	48.04
デマ 6981RT	4675	66.97	6357	91.06	5477	78.46

4.3.3 仮説 3. 「行動を促す」内容であるツイートは RT されやすい

情報発信系 RT 内において，被験者 5 名以上が選択した「行動を促す」ツイートが 19 件あった．その内，RT 数が 100RT を超えるものは 14 件あった．19 件の平均は 233RT で，多いものでは 700RT を超えるものもあった．この結果から，仮説 3 は正しいと言える．

4.3.4 仮説 4. 「ネガティブな」内容であるツイートは RT されやすい

情報発信系 RT 内において，被験者 5 名以上が選択した「ネガティブな」ツイートが 27 件あった．その内，RT 数が 100RT を超えるものは 19 件あった．27 件の平均は 251RT で，多いものでは 1000RT 近いものもあった．この結果から，仮説 4 は正しいと言える．

4.3.5 仮説 5. 「不安を煽る」内容であるツイートは RT されやすい

情報発信系 RT 内において、被験者 5 名以上が選択した「不安を煽る」ツイートが 46 件あった。その内、RT 数が 100RT を超えるものが 36 件あった。46 件の平均は 258RT で、多いものでは 800RT を超えるものもあった。この結果から、仮説 5 は正しいと言える。

4.4 考察

上記の結果より、我々の立てた 5 つの仮説は正しく、デマは「行動を促す」「ネガティブな」「不安を煽る」内容が多く、また「行動を促す」「ネガティブな」「不安を煽る」ツイートは RT されやすいことがわかった。つまり、デマは拡散されやすいと言える。

5. まとめと今後の課題

本稿では、デマとデマ訂正 RT の特徴を明らかにし、仮説の証明を行うために、震災後に RT されたツイートの解析を行い、デマは拡散されやすいことを証明した。RT の分析の中で、時間経過と共に RT の分類に変化がみられる場合を確認できた。例えば、当初はジョークとして拡散された情報が情報発信系に変化し、その結果デマとなってしまうのは、特に深刻である。そのため、情報発信系だけでなく、他の種類についても観察を行う必要がある。また、災害時におけるデマと通常時におけるデマを比較し、災害時特有の特徴についても抽出したい。

今後は、デマの訂正は拡散されにくいことを証明するとともに、デマ訂正 RT の自動検出を行い、ユーザに提示するシステムの実装を行う予定である。

参考文献

- 1) M. Sakaki, Y. Okazaki and Y. Matsuo. Earthquake shakes Twitter users: real-time event detection by social sensors, in Proc. international conference on World Wide Web (WWW), 2010.
- 2) E. Aramaki, S. Maskawa and M. Morita. Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter, in Proc. Conf. on Empirical Methods in Natural Language Processing (EMNLP), 2011.
- 3) MJ. Paul and M. Dredze. You Are What You Tweet: Analyzing Twitter for Public Health. To appear in Proc. Fifth Int'l. AAAI Conf. on Weblogs and Social Media (ICWSM), 2011.
- 4) Back, M. D., Küfner, A. C. P., & Egloff, B. (2010). The emotional timeline of September 11, 2001. Psychological Science, 21(10), 1417-1419.
- 5) Cohn, M.A., Mehl, M.R., & Pennebaker, J.W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. Psychological Science, 15, 687-693

- 6) S. Vieweg, AL. Hughes, K. Starbird and L. Palen. Microblogging during two natural hazards events: what Twitter may contribute to situational awareness. In Proc. Conf. on Human Factors in Computing Systems (CHI), pp. 1079-1088, 2010
- 7) Y. Qu, PF. Wu and X. Wang. Online Community Response to Major Disaster: A Study of Tianya Forum in the 2008 Sichuan Earthquake. Proc. Hawaii Int'l. Conf. on System Sciences (HICCS), 2009.
- 8) Twitter ブログ「リツイート機能を公開しました」
http://blog.jp.twitter.com/2010/01/blog-post_8440.html
- 9) 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, セレクション社会心理学—16,pp.45-49,(1997)
- 10) 林幸雄 (2007)『噂の拡がり方 ネットワーク科学で世界を読み解く』化学同人
- 11) TwiDema::Twitter のデマ検索サービス
<http://twidema.info/>
- 12) 松永英明 公式サイト 絵文録ことのは「震災後のデマ 80 件を分類整理して見えてきたパニック時の社会心理」
<http://www.kotono8.com/2011/04/08dema.html>
- 13) 荻上式 BLOG「東北地方太平洋沖地震, ネット上でのデマまとめ」
<http://d.hatena.ne.jp/seijotcp/20110312/p1>
- 14) ついのすみか「東北地方太平洋沖地震のデマ情報まとめ」
<http://tsuinsumika.iku4.com/Entry/67/>