

東日本大震災時のツイート分析によるデマ判別に有用な特徴抽出
Extraction of Effective Feature for False Rumor Detection by Analysis of Tweets
at the Time of Great East Japan Earthquake

榎本 光†
Hikaru Enomoto

内田 理††
Osamu Uchida

鳥海 不二夫‡
Fujio Toriumi

1. はじめに

近年, Facebook や Twitter など, SNS の利用者が増加している. なかでもツイートと呼ばれる 140 字以内の短文を投稿できる Twitter は, 東日本大震災の際に速報性の高い情報の発信と収集が行われ, 大きな注目を集めた[1]. Twitter にはリツイートと呼ばれるツイートを拡散する機能があり, 情報が瞬時に拡散する. Twitter では簡単にリツイートを行ってしまうため, ツイートの内容が正しいものであるかを確認せずにリツイートをしてしまうユーザも多く, デマが拡散されやすい状態となっている.

例えば, 東日本大震災の際には以下のようなデマが拡散された.

「給油タンクの火災で酸性雨が降る」

「うがい薬を飲むと放射線に効く」

こうしたデマは人間の安全や名誉に関わるものも多いため, 迅速に発見しその存在を明らかにすることが重要である. そこで本研究では, 東日本大震災時のツイートを分析することにより, デマ判別に有用な特徴の抽出を試みた.

2. 関連研究

Twitter では, デマが一度拡散された後に, 知識あるユーザによって拡散されたデマを訂正するという特徴がある. 鳥海ら[2]は, このデマを訂正するツイート (以下, 訂正ツイートと呼ぶ) に注目することにより, デマツイートを抽出する WEB アプリケーションを作成し, 精度の評価を行なった. 梅島ら[3]は, デマツイートに見られる傾向を分析するため, 東日本大震災時に多くのユーザによって拡散されたツイートをを用いて, デマツイートの特徴を調べている. しかし, これらの研究ではツイートの時系列による変化の分析は行われておらず, ユーザ情報にも注目していない.

本研究では, 東日本大震災時のツイートの時間変化を分析し, デマ判別に有用な特徴の抽出を試みる. また, 投稿された本文だけでなく, 投稿したユーザの情報についても分析を行う.

3. ツイートの分析による特徴抽出

東日本大震災時に収集したツイートをを用いて, デマ判別に有用な特徴を抽出する. 本研究では, URL の出現頻度とユーザ情報に注目し分析を行った. 使用したデータは, 2011 年 3 月 7 日 0 時から 3 月 23 日 24 時までの期間に収集された約 4 億の日本語ツイートである (TwitterAPI の制限などから, 該当期間の日本語ツイートの 8 割程度を収集したものである). データはツイート本文の他に, 投

† 東海大学大学院工学研究科情報理工学専攻

†† 東海大学情報理工学部情報科学科

‡ 東京大学大学院工学系研究科システム創成学専攻

稿者名, Tweet 時刻, 投稿アプリ名等が保存されている.

3.1 ツイートデータの時系列分析

3.1.1 分析方法

対象データから震災時に多く拡散されたツイートを選択し分析対象とする. 選択したツイートの関連ツイートを収集するためのキーワードを設定し, 関連ツイートを抽出する. 抽出された関連ツイートの 1 時間毎のツイート数, URL の出現数を集計し, デマツイートと非デマツイートの違い, およびデマ発生前後の変化を調べる.

3.1.2 分析結果

図 1, 2 にデマツイート (コスモ石油, うがい薬) の分析結果の例, 図 3, 4 に非デマツイート (ヤシマ作戦, ソフトバンク) の分析結果の例を示す. デマツイートにはデマ発生後にツイート数だけ増加し, 一定期間後, URL を含むツイート数も増加するという特徴が多く見られた. 一方, 非デマツイートでは, 同じような特徴は見られなかった. そのため, この特徴はデマ判別に有用な特徴の一つであると考えられる. 今回は URL のリンク先の情報には注目しなかったが, URL のリンク先の情報を分析することで, よりデマツイートの判定に有用な特徴が抽出できると考えられる.

3.2 ユーザ情報による分析

3.2.1 分析方法

デマを拡散したユーザ, 訂正ツイートを拡散したユーザ, その他のユーザ (一般ユーザ) の 3 種類にユーザを分類し, それぞれの分類ごとにツイート数, フォロワー数, フォロワー数について分析した. ユーザ情報の取得には TwitterAPI を使用した. 震災当時のユーザ情報を取得することが困難であったため, 2013 年 6 月時点での値を用いている. 一般ユーザのツイートは震災発生日の 3 月 11 日のツイートより取得した.

デマツイートおよび訂正ツイートは以下の手順で収集を行った.

- (1) 実際に拡散されたデマツイートを手動で選択する.
- (2) 関連するツイートを収集するためのキーワードを設定する.
- (3) デマおよび訂正ツイートを収集するためのキーワードを設定する.
- (4) (2) で設定したキーワードを含むツイートを全てのツイートデータから抽出する.
- (5) (4) で収集したツイートから, (3) で設定したキーワードを含むツイートを抽出する.
- (6) (5) で収集したツイートから, 否定語を含むツイートを抽出し, 抽出したツイートを訂正ツイート, 抽出された後に残ったツイートをデマツイートとする.

収集したツイートのキーワード設定例を表 1 に示す.

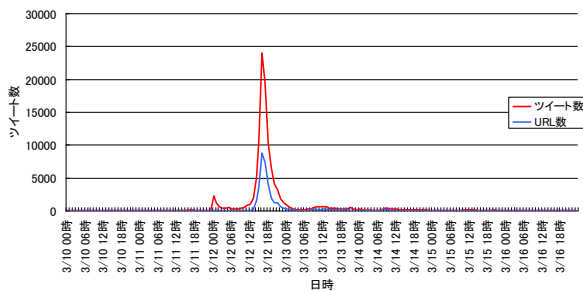


図1 コスモ石油に関するツイート数と URL 出現数

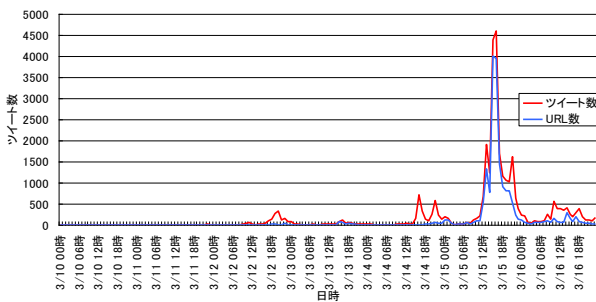


図2 うがい薬に関するツイート数と URL 出現数

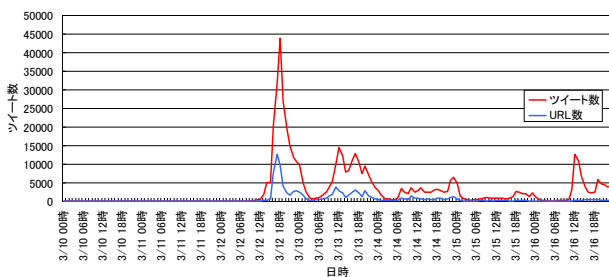


図3 ヤシマ作戦に関するツイート数と URL の出現数

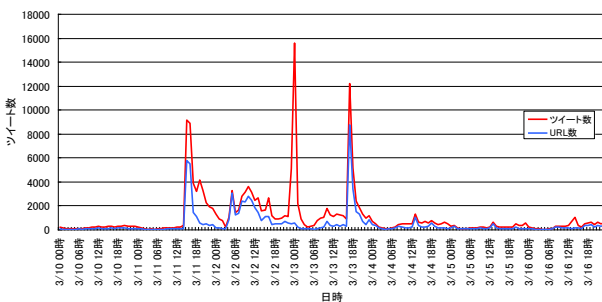


図4 ソフトバンクに関するツイート数と URL の出現数

3.2.2 分析結果

各 5000 ユーザの分析結果を表 2~4 に示す。これらの結果より、一般ユーザに比べて、デマを拡散したユーザと訂正ツイートを拡散したユーザはツイート数、フォロー数、フォロワー数が共に少ないことが分かった。しかし、デマを拡散したユーザと訂正ツイートを拡散したユーザについてはあまり差が見られなかった。

表1 各キーワードの設定 (コスモ石油)

選択したツイート	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
関連ツイート収集キーワード	コスモ石油
デマツイート収集キーワード	有害物質, 酸性雨
否定語	デマ, ガセ, 嘘, うそ, 否定, 誤報

表2 一般ユーザの情報

	ツイート数	フォロー数	フォロワー数
平均	44592.1	715.2	1268.2
中央値	17779.5	218.0	240.0
標準偏差	88157.1	4530.2	11516.4

表3 デマを拡散したユーザの情報

	ツイート数	フォロー数	フォロワー数
平均	10602.0	414.4	1357.6
中央値	5139.0	159.0	136.0
標準偏差	14713.4	3095.0	38969.6

表4 訂正ツイートを拡散したユーザの情報

	ツイート数	フォロー数	フォロワー数
平均	12217.6	643.1	866.3
中央値	5551.5	198.5	172.0
標準偏差	20252.3	5030.1	5912.2

4. まとめと今後の課題

本研究では、東日本大震災時のツイートを URL の出現頻度、ユーザ情報に注目して分析を行い、デマ判別に有用な特徴抽出を試みた。

今後の課題としては、出現する URL のリンク先にどのような特徴があるか、デマを拡散したユーザが訂正ツイートを拡散する割合がどの程度であるか等の調査が挙げられる。また、これらの特徴を用いて、デマツイートを自動判別するシステムの構築を目指していきたい。

謝辞

本研究を行なうにあたり、ツイートデータの収集に協力していただいたクックパッド株式会社の兼山元太氏に感謝する。

参考文献

- [1] 鳥海不二夫, 篠田孝祐, 栗原聡, 榎剛史, 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.
- [2] 鳥海不二夫, 篠田孝祐, 兼山元太, “ソーシャルメディアを用いたデマ判定システムの判定精度評価”, 情報処理学会デジタルプラクティス, Vol.3, No.3, pp.201-208, 2012.
- [3] 梅島彩奈, 宮部真衣, 荒牧英治, 灘本明代, “災害時 Twitter におけるデマとデマ訂正 RT の傾向”, 情報処理学会研究報告, Vol.2011-IFAT-103, No.4, pp1-6, 2011.