

マイクロブログではどの程度流言が訂正されるのか？ － Twitter 上での1年間の流言訂正情報の分析－

How Much Rumor-Correction Information is Posted on Microblog? Analysis of One Year's Rumor-Correction Information on Twitter

宮部真衣[†]
Mai Miyabe

灘本明代[‡]
Akiyo Nadamoto

荒牧英治^{†,§}
Eiji Aramaki

1. まえがき

近年, Twitter¹⁾などのマイクロブログが急速に普及している. 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 [1, 2, 3, 4]. 一方で, マイクロブログは情報漏洩や流言の拡散などの問題も抱えている. 東日本大震災においても, 様々な流言が拡散された [5].

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, デマといった概念がある. 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 (悪意をもった捏造か自然発生か) は問わないものとする. よって, 最終的に正しい情報であっても, 発言したときに十分な根拠がない場合は, 流言とみなす.

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, 誤った行動を取らせたりするなど, 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.

これまでに我々は, 人間によって投稿された流言の訂正 (以降, 訂正情報) をもとに, 自動的に流言情報を収集するサービス“流言情報クラウド”を構築してきた [6]. 本稿では, 流言情報クラウドにより1年間収集した流言情報をもとに, 人間によって訂正される流言情報について分析した結果を報告する.

2. 分析の概要

本稿では, 流言情報クラウドにより1年間収集したデータを分析する. 本章では, 流言情報クラウドについて述べた後, 対象とするデータセットについて述べる.

2.1 流言情報クラウド

流言情報クラウドは, リアルタイムに流言情報を蓄積し, 蓄積した情報をユーザに提供することにより流言の拡散防止を目指したサービスである.

流言情報を蓄積するためには, ある情報に流言が含まれているかを判定する必要があるが, 人間が信じてしまうような流言を自動的に流言だと判定することは極めて難しい. そこで, 本研究では, ある情報が流言かどうかを判定し, 流言情報を直接収集するのではなく, 訂正情報を抽出することにより, 間接的に流言情報を収集する.

[†]京都大学デザイン学ユニット

[‡]甲南大学知能情報学部

[§]科学技術振興機構さきがけ

¹⁾<http://twitter.com/>

表 1: 訂正キーワード

デマ, 嘘, ツリ, 釣り, 偽情報, ガセ, ネタ, 誤報, 都市伝説, ウソ, 狂言, 迷信, 間違い, いたずら, チェーンメール
--

表 2: 流言抽出のパターン

	PTN1	PTN2
<PTN1> などという* <PTN2>		
<PTN1> という* <PTN2>		
<PTN1> っていう* <PTN2>		
<PTN1> って* <PTN2>	「*」	訂正キーワード (表 1)
<PTN1> っていう話は <PTN2>	『*』	
<PTN1> は* <PTN2>	“*”	
<PTN1> という根拠のない		
<PTN1> という根拠薄弱な		
<PTN1> といった根拠なし		

<PTN1> <PTN2> には該当する文字列が入る.

*はワイルドカードを示す.

流言情報クラウドは, 以下の4つの機能により構成される.

- (1) クローリング機能: Twitter から訂正キーワード (表 1) を含むテキストを収集する
- (2) 訂正情報判定機能: 我々の提案する訂正情報分類器を用いて, テキストが訂正情報かどうかを判定する
- (3) 流言情報抽出機能: 訂正情報に含まれる流言情報を抽出する
- (4) 流言情報管理機能: 流言情報データベースの管理 (検索, 登録, 修正) を行う

(3) 流言情報抽出機能では, 収集した訂正情報のうち, 明示的に情報の真偽を言及しているものから流言情報を収集する. 流言の抽出手順を以下に示す.

1. クローリング機能によって収集したツイートから, 訂正情報判定機能により訂正情報を抽出する.
2. 手順 1 で抽出した訂正情報から, “「～」というデマ”というパターン (表 2) を用いて流言内容 (“～”部) を抽出する.

2.2 対象データセット

本研究では, 流言情報クラウド上で流言情報の収集を開始した 2012 年 6 月 22 日から 1 年間のデータを用いることとした²⁾.

²⁾なお, 収集期間中には Twitter API の仕様変更によって, データの収集ができなかった日も存在する.

表3: 同じ流言の表現のバリエーションの例

- iPhone を電子レンジでチンすると直ぐに充電ができる
- iPhone を電子レンジでチンすると急速充電できる
- iPhone を電子レンジでチンすると充電できる
- iPhone が電子レンジで充電できる

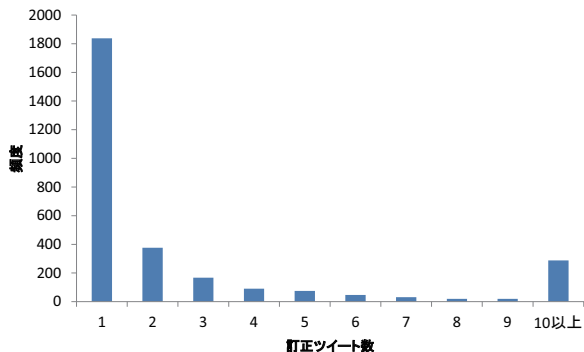


図1: 訂正ツイート数の分布

表4: 1日あたりの訂正ツイート数, 訂正された流言数

	訂正ツイート数 (件)	訂正された流言数 (件)
平均	109.4	14.0
標準偏差	278.9	5.3
最大	4507.0	34.0
最小	0.0	0.0

まず, 2012年6月22日~2013年6月21日までに訂正キーワードをもとに収集・抽出した訂正ツイートから, 2.1節で述べた手順により, 人間によって訂正された流言情報を抽出した. なお, この手順では同じ流言の異なる表現のバリエーション(表3)も抽出される. 同一の流言を指し示すと考えられる異なる表現の流言は, 1日毎あるいは1年全体で1つにまとめて分析を行う.

3. 分析結果

3.1 流言別の訂正数

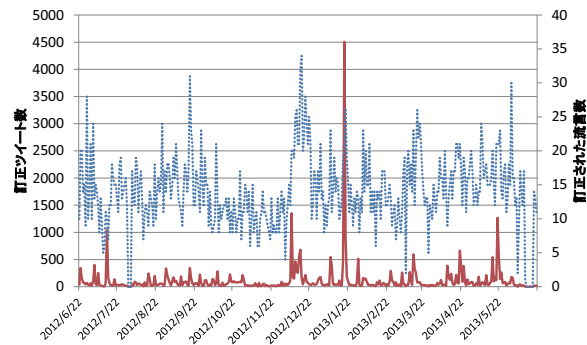
1年全体で訂正された流言数について確認したところ, 本収集期間においては, 2953件の流言に対する訂正情報が投稿されていた. これらの流言に対して, どの程度の訂正ツイートが投稿されたのかの分布を図1に示す.

図1より, 訂正ツイート数が1件のみしか投稿されない流言が多く, 全体の62%を占めている. 一方, 10件以上訂正ツイートが投稿された流言は全体の10%程度(288件)であった. ただし, 訂正ツイート数と実際に拡散された流言ツイート数の関係については未だ明らかではないため, 今後流言ツイートについての分析を行い, 流言拡散防止における訂正ツイート数の少ない流言の取り扱い方針について検討する必要がある.

3.2 1年間の訂正数推移

表4に1日あたりの訂正ツイート数および訂正された流言数を示す. また, 図2に1年間の訂正ツイート数, 訂正された流言数の推移を示す.

1日あたりに訂正される流言数は, 平均14件, 最大34件であった. 一方, 1日あたりの訂正ツイート数は平均109件, 最大4507件であった. 図2より訂正ツイート数



実線が訂正ツイート数, 点線が訂正された流言数の推移を示す.

図2: 訂正ツイート数, 訂正された流言数の推移

の急激な増加がみられる部分があるが, この際, 必ずしも訂正される流言数の増加は見られない. これは, 多数のユーザによる特定の流言に対する訂正ツイートが投稿されたことにより, 訂正ツイート数のみが急激に増加している. 例えば, 図2の2013年1月19日を見ると, 流言数自体は20件と大きな増加は見られないが, 「iPhoneを電子レンジでチンすると直ぐに充電ができる」という特定の流言に対する訂正4477件が発生したことにより, 訂正ツイート数の急増が見られた.

4. むすび

本稿では, 人間による流言訂正情報に基づき自動的に流言情報を収集するサービス“流言情報クラウド”によって1年間流言情報を収集し, Twitter上で訂正される流言情報の推移について分析を行った.

1年分のデータを分析した結果, 1年間に訂正された流言の大部分は少数のユーザのみが訂正したものであり, 10件以上の訂正ツイートが投稿された流言は10%程度であることが観察された. また, 1日あたりに訂正される流言数は平均14件(標準偏差5.3)程度であるが, 多数のユーザが訂正する特定の流言の出現により, 訂正ツイート数が急激に増加する場面も観察された.

今後は, 流言情報自体の投稿傾向と併せて分析することにより, 効率的な流言拡散防止手法の検討を行う.

謝辞

本研究は, JST 戦略的創造研究推進事業の助成を受けた.

参考文献

- [1] 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.
- [2] 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.
- [3] 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.
- [4] 宮部真衣, 荒牧英治, 三浦麻子. 東日本大震災における twitter の利用傾向の分析. 情報処理学会研究報告. GN, [グループウェアとネットワークサービス], Vol. 2011, No. 17, pp. 1–7, 2011.
- [5] 荻上チキ. 検証東日本大震災の流言・デマ. 光文社, 2011.
- [6] 宮部真衣, 梅島彩奈, 荒牧英治, 灘本明代. 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. マルチメディア, 分散, 協調とモバイル (DICOMO2012) シンポジウム, pp. 1442–1449, 2012.