

## 情報の信頼性判断のための流言テキストの特徴分析

柿本 大輔<sup>†</sup>荒牧 英治<sup>††</sup>宮部 真衣<sup>†</sup><sup>†</sup>和歌山大学システム工学部<sup>††</sup>奈良先端科学技術大学院大学研究推進機構

### 1 はじめに

近年、SNSの普及により、ネットワーク上でのコミュニケーションが活性化している<sup>\*1</sup>。特にマイクロブログにおいて、ユーザはリアルタイムに多種多様な情報を取得・発信することが可能である。一方、その手軽さから、有用な情報だけでなく、流言も伝播されやすい<sup>\*2</sup>。災害時などでは、流言の伝播によってユーザが誤った情報を取得し、行動することで、深刻な問題が引き起こされる可能性がある。そのため、流言の拡散を防止する仕組みが必要である。

流言の拡散防止においては、ユーザによる情報の信頼性判断が不可欠である。しかし、情報の信頼性判断基準は各ユーザで異なり、誰もが正確な判断を行うことが出来るとは限らない。本研究では、流言である可能性が高い情報に含まれる、特徴的なテキスト表現について分析する。

### 2 関連研究

流言に関する研究は多方面からなされている。特にTwitter<sup>\*3</sup>などのマイクロブログ特有の性質に着目して、流言伝播の対策を試みる研究も数多く見られる。

荒川らの研究[1]ではTwitterにおけるリツイート機能に着目し、4種類の特徴量を用いて分類実験を行い、リツイートされやすいツイートの特徴を検討した。また、中原らの研究[2]では、ツイートに対するリツイート回数及び非公式リツイートに付加されたコメントを用いて、ツイートがデマであるかどうかの判断を支援する手法を提案した。これらの研究では、リツイート機能から、Twitterにおける情報伝播を分析し、流言の拡散防止を試みている。本研究では、リツイートなどのTwitter特有の情報ではなく、テキストの文型的表現の違いによる、信頼性判断支援手法を検討する。

福島らはWebページの情報を対象に、Webサイトの信頼性を決定付ける尺度を用いて、情報の信頼性を自動推定する手法を提案した[3]。この研究では、情報の

真偽に関わらず、ユーザが正しいと認識してしまう情報を、信頼性のある情報と定義している。ただし、この手法はWebページを対象としており、マイクロブログに対して適用ができるとは限らない。そこで本研究では、マイクロブログのような短文の情報にも適用可能な信頼性判断について検討する。

### 3 情報の信頼性判断

#### 3.1 信頼性の定義

本研究では、福島らの研究[3]と同様に「情報の真偽に関わらず、ユーザが正しいと認識してしまうもの」を信頼性のある情報と定義する。この定義に基づき、Twitterなどのマイクロブログ及びWebページなどに特有の情報ではなく、日本語の文型としてのテキスト表現に着目して分析を行う。

#### 3.2 テキストと信頼性に関する予備調査

これまでに我々は、東日本大震災直後のTwitterを材料とし、どのような流言が深刻な影響を与えるのかを、主観評価及び修辞ユニット分析により分析した[4]。この研究では、流言の内容の有用性、有害性、ネガティブさ、行動促進、不安扇動、尤もらしさ（本稿における信頼性）、伝聞情報の7項目について、7名の実験協力者により主観評価を行った。この評価結果から、まず、情報の信頼性に影響し得る項目について検討する。

尤もらしさ（信頼性）と、その他の6項目との相関について確認した結果（尤もらしさと有害性の相関係数は0.233、尤もらしさと伝聞情報の相関係数は0.231）、有害性及び伝聞情報のみ、尤もらしさととの弱い相関が見られた。この2項目の内、情報の有害性は、ユーザにより判断基準が異なると考えられ、有害性の有無を明確に定義することは容易ではない。一方、伝聞情報は、伝聞を意図するテキスト表現が現れると考えられ、伝聞情報かどうかは定義しやすい。そこで本稿では、伝聞情報に着目し、信頼性判断への影響について調査する。

### 4 評価実験

#### 4.1 実験概要

本研究では、「テキスト内の伝聞表現が、ユーザの情報の信頼性判断に影響する」という仮説を立て、検証を行う。本実験では、大学生20人による情報の信頼性に関する主観評価実験を行った。主観評価では、提示

Feature Analysis of Rumor Tweets for Supporting Reliability Judgement in its Information

Daisuke KAKIMOTO<sup>†</sup> Eiji ARAMAKI<sup>††</sup> Mai MIYABE<sup>†</sup>

<sup>†</sup>Faculty of Systems Engineering, Wakayama University

<sup>††</sup>Center for Frontier Science and Technology, Nara Institute of Science and Technology

<sup>\*1</sup><http://ictr.co.jp/report/20150729000088-2.html>

<sup>\*2</sup>[http://www.google.org/crisisresponse/kiroku311/chapter\\_18.html](http://www.google.org/crisisresponse/kiroku311/chapter_18.html)

<sup>\*3</sup><http://twitter.com/>

表 1: 用いた伝聞表現とトリビアへの付加例

伝聞表現パターン	テキストの例
原文	うなぎの血液は毒である。
1. 普通形+そうです	うなぎの血液は毒だそうです。
2. 普通形+と(聞いた)	うなぎの血液は毒だと聞いた。
3. 普通形+らしいです	うなぎの血液は毒らしい。

表 2: 原文を基準とした各伝聞表現に対する評価結果の差 (情報が正しい場合)

	-2 以下 (文)	-1 (文)	0 (文)	1 (文)	2 以上 (文)
伝聞表現 1	9	44	91	26	8
伝聞表現 2	10	45	86	34	5
伝聞表現 3	15	41	80	37	7

されたテキスト情報に信頼性があるか、という問いに対し、5段階評価(1:強く同意しない, 2:同意しない, 3:どちらともいえない, 4:同意する, 5:強く同意する)で回答してもらった形式をとった。

評価テキストとして、Web から収集した 360 件のトリビア<sup>\*1</sup>(半数は流言(誤りを含む情報)となるよう、著者が意図的に内容を変更したもの)を利用し、原文及び書き換え文(テキストの文末に伝聞表現を付け加えたもの)を用いる。テキストの文末に付加する伝聞表現は、日本語表現文型に関する文献 [5] に記されている、代表的な 3 種類の伝聞表現を用いた。用いた伝聞表現とトリビアへの付加例を表 1 に示す。

#### 4.2 結果と考察

今回は、360 件のテキストにおける、表現ごとの実験協力者間の評価結果の中央値を代表値とし、議論する。原文を基準とした、各伝聞表現に対する評価結果の差(情報が正しい場合)を表 2 に示す。同様に、情報が誤っている場合の結果を表 3 に示す。情報の真偽に関わらず、伝聞表現が付加されたテキストの評価結果が、原文の評価結果より高く示されたテキスト数は、伝聞表現 1 で 91 件(正しい情報 34 件, 誤った情報 57 件)、伝聞表現 2 で 78 件(正しい情報 39 件, 誤った情

<sup>\*1</sup> 雑学的な知識、豆知識を指す。本研究ではユーザの事前知識及び経験に依らず、テキスト表現のみによる情報の信頼性判断について分析するため、評価テキストとしてトリビアを用いた。引用元 URL は以下に示す。

- [http://www.ideaxidea.com/archives/2011/01/animal\\_trivias.html](http://www.ideaxidea.com/archives/2011/01/animal_trivias.html)
- <http://www.madameriri.com/2011/10/20/世界の雑学うんちく/>
- [http://www.popxpop.com/archives/2007/01/post\\_89.html](http://www.popxpop.com/archives/2007/01/post_89.html)
- <http://fleshwords.web.fc2.com/hitozatsu/hitozatsu3.html>
- <http://matome.naver.jp/odai/2128082323158302801>
- <http://www.popxpop.com/archives/2007/02/55.html>
- <http://www.geocities.jp/f9305710/zatsugaku.html>
- <http://www5.plala.or.jp/zatsugaku/toribia.html>
- <http://www.zatugaku-king.com/list.html>
- <http://takech.fc2web.com/hee.html>
- <http://www.trivia-blog.com/>
- <http://netao.biz>

表 3: 原文を基準とした各伝聞表現に対する評価結果の差 (情報が誤っている場合)

	-2 以下 (文)	-1 (文)	0 (文)	1 (文)	2 以上 (文)
伝聞表現 1	4	40	77	51	6
伝聞表現 2	5	50	86	37	2
伝聞表現 3	4	37	95	39	5

報 39 件)、伝聞表現 3 で 88 件(正しい情報 44 件, 誤った情報 44 件)であった。また、伝聞表現 1 では情報が誤っている場合、情報が正しい場合と比較して、原文に対し評価結果が高いテキスト数が、23 件多かった。

表 2, 表 3 より、情報の信頼性において、伝聞表現の有無による評価結果への大きな影響は見られなかった。ユーザは情報の信頼性判断の際、伝聞表現の有無ではなく、その内容から判断することが多いと考えられる。

一方、情報が正しい場合と比較して、情報が誤っている場合、伝聞表現 1 では、原文よりも評価結果が高いテキストの割合が多い傾向が見られており、伝聞表現 1 は、他の伝聞表現よりも情報の信頼性に影響を与える可能性がある。

#### 5 おわりに

本研究では、情報の信頼性判断支援を行うために、流言である可能性が高い情報に含まれる、特徴的なテキスト表現について分析した。予備調査結果に基づき、「伝聞表現は、ユーザの情報の信頼性判断に影響を与える」という仮説を立て、実験協力者による情報の信頼性に関する主観評価を行った。

評価の結果、伝聞表現による情報の信頼性への、大きな影響は見られなかった。しかし、情報が誤っている場合、伝聞表現 1 は他の伝聞表現と比較して、信頼性に影響を与える可能性があることがわかった。

今後は、伝聞表現以外の文章表現による情報の信頼性判断への影響を検討する。また、トリビア以外の、より一般的な内容のテキストを対象とした評価実験を行う。

#### 謝辞

本研究は、JSPS 科研費 15H05317 の助成による。

#### 参考文献

- [1] 荒川唯, 亀田堯宙, 相澤彰子, 鈴木崇史: Retweet に着目した広がりやすい Tweet の特徴分析, 第 74 回全国大会講演論文集, Vol.2012, No.1, pp. 617-618(2012).
- [2] 中原英美, 富永一成, 牛尾剛聡: リツイート構造を用いたデマ拡散防止支援手法, DEIM Forum, F2-3(2012).
- [3] 福島隆寛, 内海彰: Web ページの信頼性の自動推定, 知能と情報(日本知能情報フレンジイ会誌), Vol19, No.3, pp. 239-249(2007).
- [4] 宮部真衣, 田中弥生, 西畑祥, 灘本明代, 荒牧英治: マイクロブログにおける流言の影響の分析, 自然言語処理, Vol. 20, No. 3, pp. 485-511(2013).
- [5] 友松悦子, 宮本淳, 和栗雅子: 改訂版 どんときどう使う日本語表現文型 200, pp. 100-144, アルク (2013).