

## 訂正投稿の傾向を利用したデマ訂正ツイートの抽出 Extractions of Rumor Correction Tweets Based on Characteristic of Their Posting

渡邊 建太<sup>†</sup>  
Kenta Watanabe

山田剛一<sup>†</sup>  
Koichi Yamada

絹川博之<sup>†</sup>  
Hiroshi Kinukawa

### 1. はじめに

Twitter は 140 文字以内の文章を投稿 (以下, ツイート) し, それを共有できるマイクロブログサービスである. 速報性と手軽さから, Twitter には多くのユーザが様々な内容をツイートしている. 有益なツイートも多く投稿されているが, それと共にデマが拡散されるという問題も起きている. デマを拡散するユーザの多くが, その情報は正しいと思って拡散している [1]. よって, 事前にデマであることを確認できるようにすることにより, デマ拡散を防止できる可能性があると考えられる.

デマの内容を含むツイート (以下, デマツイート) は多種多様であるため, 直接デマツイートを検出し収集することは困難である.

しかしながら, その情報がデマであることを指摘しているツイート (以下, デマ訂正ツイート) では「デマ」などの訂正表現を含むものが多く, 検出が可能である [1].

また, デマ訂正ツイートを投稿したユーザ (以下, デマ訂正投稿者) は他にもデマ訂正ツイートを投稿していることがある.

これらの特性に着目して, 投稿者とその投稿表現の特徴を利用したデマ訂正ツイートの抽出方法を提案する.

## 2. デマ訂正ツイートとデマ訂正投稿者の傾向

### 2.1 デマ訂正ツイートとは

デマツイートとは, 情報の内容に誤りがあるツイートのことを言い, 1 章で述べた通りデマツイートは内容が多種多様であり, 抽出することが困難である. 一方, デマ訂正ツイートには訂正を表す特定の語や表現が含まれていることが多い. また, 訂正を表す語や表現は比較的抽出が容易であると考えられる. 本研究におけるデマ訂正ツイートは, 情報の内容の誤りを指摘しているツイートとする.

ただし, 以下のツイートはデマ訂正ツイートではないものとする.

- (1) 単に「デマ」とだけ言って, デマの訂正内容を含まないツイート
- (2) 個人的なことや会話の中で「デマ」と表現しているツイート
- (3) リンクしている記事の中でデマを指摘しているツイート

なお, (3) の Web の記事を紹介しているツイートにユーザ自身がデマを指摘しているコメントがあるツイートはデマ訂正ツイートとする.

また, Twitter のデマに関する調査を行った結果, 情報が訂正されているツイートにはデマ訂正に関するキーワードが含まれている場合が多く見られた.

そこで本研究では, デマ訂正に関するキーワードのうち, 表 1 に示す語 (以下, 訂正語) のいずれかを含むツイートを抽出する. また, 渡邊ら [2] の研究から訂正語の前後の特徴的な表現の有無はデマ訂正ツイート抽出に役立つことが期待できるので, この特徴を利用してデマ訂正ツイートの抽出を試みる.

表 1 対象とする訂正語一覧

対象とする訂正語	デマ, ガセ, 誤報, 噂嘘, ウソ, 偽情報

### 2.2 デマ訂正投稿者の傾向

デマ訂正投稿者の中には, 1 つではなく様々な情報の誤りを指摘しているユーザがいる.

デマ訂正ツイート自体は Twitter 全体から見れば, 数はとても少なく, 普通のツイートに埋もれてしまっている. そこで, 様々なデマを指摘しているデマ訂正投稿者に着目し, このような投稿者をリスト化することで効率的にデマ訂正ツイートを抽出することを試みる.

複数回デマ訂正ツイートを投稿したユーザ (以下, 有益なデマ訂正投稿者) を集めたリストを作成し, 彼らの投稿したツイート群を利用して実験を行う.

## 3. デマ訂正ツイート抽出システム

提案するデマ訂正ツイート抽出システムは 2 つのフェーズがある.

- (1) Twitter からデマ訂正ツイートを抽出し, それを複数投稿したユーザを有益なデマ訂正投稿者としてリストに登録するフェーズ
  - (2) 有益なデマ訂正投稿者リストの全てのユーザのツイートからデマ訂正ツイートを抽出するフェーズ
- 以下では, 2 つのフェーズについて述べる.

### 3.1 有益なデマ訂正投稿者の発見とリストへの登録

3 章の (1) で述べた有益なデマ訂正投稿者を発見し, 有益なデマ訂正投稿者リストに登録するシステムを図 1 に示す.

<sup>†</sup>東京電機大学大学院 未来科学研究科  
Graduate School of Science and Technology for Future Life,  
Tokyo Denki University

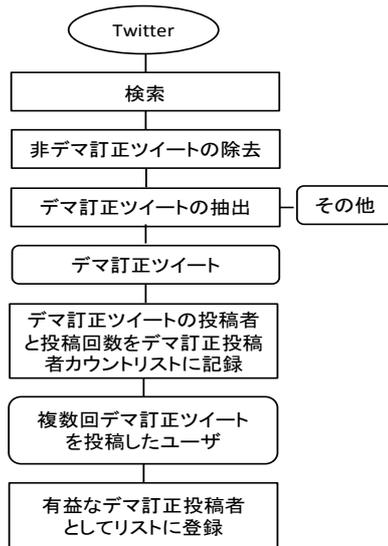


図 1 有益なデマ訂正投稿者をリストに登録するフェーズのシステム図

はじめに、Twitter 検索の機能を利用して、2.1 節で述べた訂正語のいずれかを含むツイートを取得する。なお本システムでは、Twitter4J を用いる [3]。そして、非デマ訂正ツイートを除去し、デマ訂正ツイートを抽出、収集する。同時に抽出したデマ訂正ツイートを投稿したユーザーと投稿回数をデマ訂正投稿者カウントリストに登録する。そして、複数回デマ訂正ツイートを投稿したユーザーがいれば、そのユーザーを有益なデマ訂正投稿者とし、有益なデマ訂正投稿者リストに登録する。

### 3.2 有益なデマ訂正投稿者リストによるデマ訂正ツイートの抽出方法

3 章で述べた (2) のフェーズのシステム図を図 2 に示す。

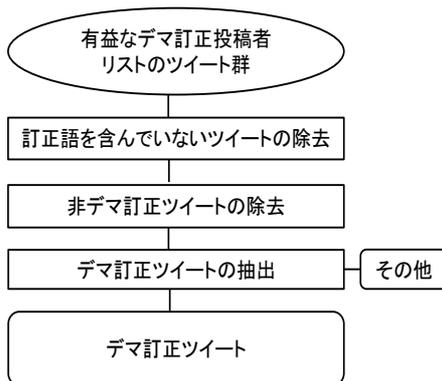


図 2 有益なデマ訂正投稿者リストのツイート群からデマ訂正ツイートを抽出するフェーズのシステム図

3.1 節のシステムで得た有益なデマ訂正投稿者リストに登録されているユーザーのツイートを利用する。はじめに、訂正語を含まないツイートを除去したあと、さらに非デマ訂正ツイートを除去する。残りのツイートからデマ訂正ツイートを抽出する。

### 3.3 非デマ訂正ツイートの除去

本研究ではデマ訂正ツイートの抽出を目的としている。そこで、デマ訂正をしていないツイート (以下、非デマ訂正ツイート) は対象外とし、除去する。非デマ訂正ツイートの条件を以下に示す。

(1) リツイートではなく、13 文字以下のツイート

これは文字数が少ないためツイートの情報が不足しており、単に“デマ”とだけ言ってデマの内容が含まれていないツイートが多いため除去する。

(2) リツイートではなく、訂正語の後に「？」を含むツイート

これはツイートした本人がその情報の事実を知らず、デマであると訂正しているわけではないので除去する。

(3) リツイートではなく、「私」、「僕」、「俺」のいずれかを含むツイート

上記のいずれかの語を含むツイートでは、そのツイート内容が私的であることが多い。もしデマ訂正された内容がデマだったとしても、多くの人に影響を及ぼす可能性が低いいため除去する。

(4) リンクしている記事の中でデマを指摘しているツイート

これはユーザー自身が訂正していないため除去する。

### 3.4 デマ訂正ツイートの抽出

3.3 節の処理から残ったツイートの中で以下の (1) または (2) の条件に該当するツイートをデマ訂正ツイートとして抽出する。○○は訂正語のいずれかを指す。

(1) 以下のいずれかを含むツイート

- ① ○○です ② ○○だった ③ ○○だそうです  
④ ○○らしい ⑤ デマに騙され ⑥ デマが出回る  
⑦ デマを流す ⑧ というのはガセ ⑨ ガセネタだった

(2) 語尾が「デマ」または「ガセ」で終わるツイート

## 4. おわりに

本研究は、デマ訂正投稿者の傾向を利用したデマ訂正ツイートの抽出方法を提案した。

今後の課題として、デマ訂正ツイートの抽出条件の充実や有益なデマ訂正投稿者リストの有無によるデマ訂正ツイートの収集効率の比較実験が挙げられる。

### 謝辞

本研究に際して、使用させていただいた Twitter4J の開発者の方々に感謝いたします。

### 参考文献

- [1] 宮部真衣, 梅島彩奈, 灘本明代, 荒牧英治. 流言情報クラウド: 人間の発信した訂正情報の抽出による流言収集. 言語処理学会第 18 回年次大会 (2012)
- [2] 渡邊研斗, 鍋島啓太, 岡崎直観, 乾健太郎. Twitter 上での誤情報と訂正情報の自動分類. 言語処理学会第 19 回年次大会 (2013)
- [3] Twitter4J, <http://twitter4j.org/ja/index.html>