

Twitter におけるデマ拡散防止支援に関する研究

A Study on Supporting to Prevent Demagogy Diffusing on Twitter

1DS08185S 中原英美 NAKAHARA Emi

1. はじめに

近年、マイクロブログの一種である Twitter の利用者が飛躍的に増加している。Twitter は、個々のユーザが「ツイート」とよばれる 140 字以内の短文を投稿し、閲覧できるコミュニケーション・サービスである。

Twitter では、リツイート (RT) と呼ばれる機能によって他のユーザの投稿を再投稿することができる。これにより、手軽に情報を広めることができる。

リツイートは情報を広めるのに有効である。しかしリツイートによって誤った情報 (デマ) が広まるのが社会的に大きな問題となっている。

このような背景の下、本研究では、リツイートによるデマの拡散防止を支援する機構を開発することを目的とする。

2. アプローチ

リツイートによるツイートの拡散は図 1 のように木構造的に行われる。したがって、基本的にはリツイートされる情報の提供者は、自分が発したツイートがリツイートによってどれだけ拡散し、どのユーザに配信されたかは解らない。

また、Twitter のフォロー関係は単方向である。すなわち、自分がフォローしたユーザを他のユーザもフォローしているとは限らないため、ツイートの情報伝播は単方向的に行われる。その結果、リツイートされた情報に対してそれを読んだユーザが別のツイートを行うなどで議論や付加情報が提供されたとしても、別の経路でリツイートを受け取ったユーザには伝達されない。

本研究では上記の問題点を解決するために、ユーザが自分の読んだリツイートの信憑性を確認したいと思ったとき、別の経路で配信されているリツイート上で投稿された訂正情報の有無などを確認可能とすることにより、ユーザのリツイートに関する信憑性の判断の支援を行う手法を提案する。

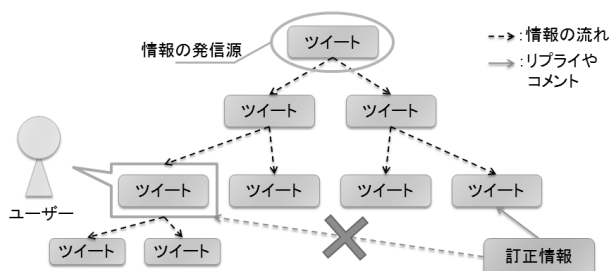


図 1: リツイートによる情報の拡散

3. 提案手法

リツイートは、公式リツイートと非公式リツイートの 2 種類に分類できる。公式リツイートは、発言元のツイートがそのまま表示されるのに対し、非公式リツイートは「コメント RT: @username 引用文」という形式のツイートとして投稿される。

我々は、非公式リツイートに付加されたコメントをリツイートしたツイートに対する意見として抽出する。抽出したコメントを利用し、リツイートが拡散していく状況で信憑性に疑問が持たれているかどうかを、ユーザに提示することで、信憑性が低い情報に関する注意を促す。

3. 1. 訂正ツイート

リツイートに対して、その情報に関して「デマ」など、その信頼性に対して否定的な意見をコメントとして付加している非公式リツイートを「訂正ツイート」と呼ぶ。訂正ツイートを取得するために、Twitter API の search メソッドを使用し、「username」+「引用文」をクエリとして検索して得られた非公式リツイートからコメントを抽出する。抽出したコメントが「デマ」「間違い」「ガセ」「釣り」「嘘」という単語を含む場合、そのコメントは引用文の内容に対する訂正情報であると考えられる。

3. 2. 危険度

あるツイートがリツイートされたうち、リツイートの中で、訂正ツイートが占める割合のことを危険度と呼ぶ。

非公式リツイート数と訂正ツイート数については上記の訂正ツイートを取得する際にカウントし、公式リツイート数は Twitter API の statuses/:id/retweeted_by メソッドを使用して取得する。取得した公式リツイート数、非公式リツイート数、訂正ツイート数を用いて危険度を算出する。

具体的には、公式リツイート数を O_n 、非公式リツイート数を U_n 、訂正ツイート数を C_n とすると、危険度は以下の式で求められる。

$$(\text{危険度}) = \frac{|C_n|}{|O_n| + |U_n|}$$

4. おわりに

本稿では、Twitter においてデマかどうかをユーザが判断することを支援するための基準として危険度・訂正ツイートを提案した。提案手法では、非公式リツイートに付加されたコメントを用いて、ユーザに提示する二つの情報、「訂正ツイート」「危険度」を取得する。それによって、リツイートが拡散していく状況で信憑性に疑問が持たれているかどうかを、ユーザが簡単に調査できるようにすることで、信憑性が低い情報に関する注意を促す。