

時系列解析による SNS 上のデマ状態推定モデル

牛込龍太郎^{†1} 松田健^{†2} 園田道夫^{†3} 趙晋輝^{†1}

概要 :

SNS(Social Networking Service)は各ユーザーが不特定多数のユーザーに対して手軽に情報発信できるため、多くの分野で活用されている。しかしながら、SNS で伝搬される情報の真偽を直ちに判断することは難しい場合もあり、誤った事実や虚偽の情報いわゆるデマが拡散されやすいことが社会問題として注目されている。SNS 上でのデマに対する世間の警戒は高まりつつあるものの、実社会へ影響を与えるようなデマの発生は根絶には至っていないため、このようなデマがどのようなタイミングで発生し、どのように拡散・収束するか調査することが本研究の目的である。本稿ではデマの実データの一部を利用し、二項分布を用いたデマ状態推定モデルを提案した。

キーワード : SNS, Twitter, 二項分布, デマ

1. 序章

SNS は不特定多数のユーザーによって膨大かつ多様な情報がやりとりされており、そのデータを目的とする SNS の活用場面は今なお広がっている。しかし多くの SNS において投稿の内容は事前に確認されることなく投稿されるため、SNS 内で誤った情報が拡散されるケースも存在し、一部は社会問題となっている[1][2]。今後もこのようないわゆる「デマ」が拡散されるケースは発生すると考えられ、SNS データの利活用を進めていく上で SNS の投稿内容が正しいかどうか判断できるようになる必要がある。そのために、デマが SNS ユーザーの間でどのように拡散されていくのかを解析していくことは有意義であると考えられる。

本研究では SNS への最も普及している SNS の一つである Twitter の実データを用いて、ユーザーがデマにどの程度騙されているかを示す二項分布を応用した状態推定モデルの検討を行った。なお、本稿における「デマ」とは「ある SNS ユーザーが SNS コミュニティあるいは現実社会を混乱させることを目的として拡散させる事実と異なる情報」と定義する。

2. 関連研究

Twitter 内で情報が拡散される動きに対してモデリングを行った研究はいくつか存在する。池田ら[3]は いわゆる「ロコミ」に着目しユーザーの生活パターンと複数の情報源からの情報発信を考慮した新たな情報拡散モデルを提案し、実際のデマ情報の再現を行うことでモデルの妥当性を確認している。佐藤ら[4]は感染症数理モデルの 1 つ SIR モデルを応用し、デマ情報に関してユーザーが 3 つの状態にあると仮定した上で、デマ情報の再現を行っている。

3. 提案手法

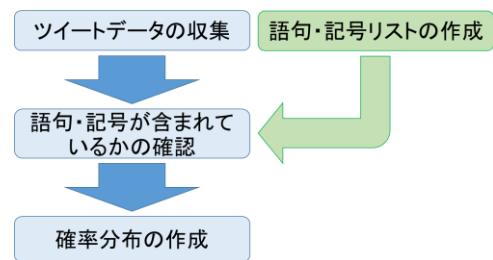


図 1 提案手法の流れ

本研究の提案手法の流れは図 1 の通りである。はじめに Twitter への投稿「ツイート」を収集する。この時ツイートデータの収集とは別に、デマに騙されているツイートに出現しやすいと考えられる語句・記号のリストを作成する。次に収集した各ツイートの文中に語句・記号リストの要素が含まれているかどうかを {0,1} 値で表していき、これを作成する確率分布への入力とする。

3.1 語句・記号リスト

本手法では、ユーザーがデマに騙されている程度を、ツイート中に含まれる語句・記号の種類数で対応させる。作成した語句・記号リストを表 1 に示す。なお、一部の語句については実際のツイート文中での表記揺れや口語的な語尾の変化などを考慮している。今回のデータセット内のツイートには、リストの語句全てを含むものはなかった。

表 1 語句・記号リスト

| | | |
|----|-----|------|
| ? | どっち | ムカつく |
| 最低 | 怒り | 騙り |
| 最悪 | やばい | 迷惑 |
| 酷い | 本当に | |

3.2 二項分布

二項分布とは Bernoulli 試行の一方の結果が出る回数ごと

^{†1} 中央大学
Chuo University
^{†2} 長崎県立大学
University of Nagasaki

^{†3} 情報通信研究機構
National Institute of Information and Communications Technology

の確率を示す確率分布であり、(1) 式のように示される[5].

$$\Pr[X = j] = {}_n C_j p^j (1-p)^{n-j} \dots \quad (1)$$

ここで、 p は各 Bernoulli 試行において一方の結果 (結果 A とする) がでる確率、 n は試行の総回数、 j は n 回の試行のうち結果 A が出る回数をそれぞれ表す。(1) 式を踏まえ、本研究では (2) 式のような二項分布を考える。

$$\Pr_{u,t}[X = j] = {}_n C_j p_{u,t}^j (1-p_{u,t})^{n-j} \dots \quad (2)$$

(2) 式において、 u はユーザー、 t はツイートが投稿された時刻、 $p_{u,t}$ はユーザーがデマに騙されている確率、 n は語句・記号リストの要素の個数、 j はツイート文中に含まれる要素の個数をそれぞれ表す。(2) 式のような二項分布を各時刻、各ユーザーに対して作成し、各時刻におけるユーザー同士の分布の積を取ることで、各時刻においてユーザーがその話題に対してどの程度騙されているかを示す確率分布の尤度分布を得ることができ、ベイズ推定における事前確率分布へ適用することが可能であると考えられる。

3.3 データセット

本研究で用いた Twitter データは、2018 年 5 月中旬に投稿されたあるツイート T_A に言及するものである。 T_A の内容は、日本国内のある飲食店が隣の大学の教職員から貸し切りの予約を受けたにも関わらず、無断で来店をキャンセルされたと主張するものであった。しかし実際は飲食店、大学両者とも架空の存在であった。収集したツイートデータは 1,504 ユーザー・2,243 件であり、各ツイートデータにはツイート文、秒単位での投稿時刻などが含まれている。

4. 結果

3.3 節で述べたデータセットに対して提案手法を適用した結果の一部を示す。図 2, 3 の各縦軸は確率密度を、各横軸はユーザーがデマに騙されている確率 p をそれぞれ表す。各色の折れ線は、緑線、黄線、赤線がそれぞれ 3.2 節の (2) 式において $j=0, 1, 2$ に対応している。図中の t は時刻を表しており、 t の番号は時刻の昇順になっている。本研究の結果に対応する確率分布の式は (3) 式の通りである。

$$\Pr_{u,t}[X = j] = {}_{11} C_j p_{u,t}^j (1-p_{u,t})^{11-j} \quad (j = 0, 1, 2) \dots \quad (3)$$

グラフの時系列変化としては、グラフの頂点がある時点まで右寄りに、すなわちユーザーがデマの話題により強く騙

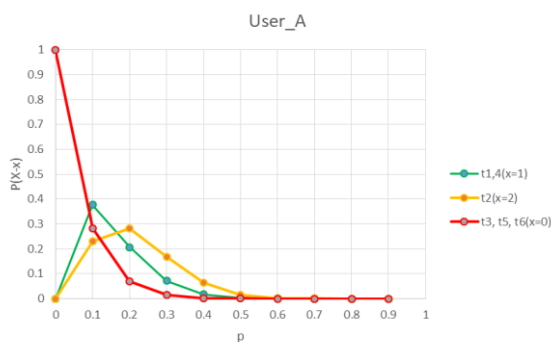


図 2 ユーザー A の確率分布

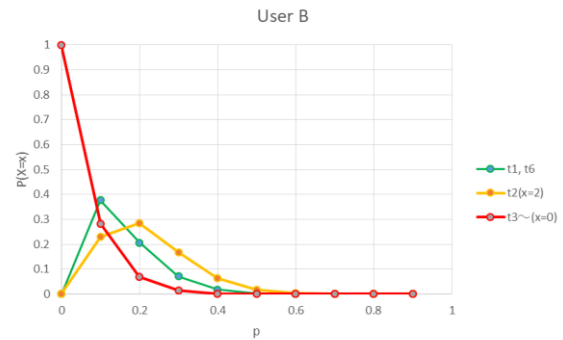


図 3 ユーザー B の確率分布

されるように変化し、その後グラフの頂点が左寄りに移動、すなわちデマの話題に騙されなくなるように変化している。ただし、デマに気づいたのちもリストの語句を含むツイートを行った場合は頂点が右寄りのグラフに一瞬移る。

5. 考察

図 2, 3 よりグラフが最大となる p がいずれの場合においても 0.3 を下回っていることが確認できる。これは本研究で使用した各ツイート中に語句・記号リストの要素が高々 2 個しか含まれていなかったためと考えられる。1 ツイートは最大 140 文字の投稿しかできないため、短い文字列においてもリストの要素を特徴量として生かせるリストの設定が必要と考えられる。またグラフの時系列変化については、騙されていたことに気づいた後も、緑線のグラフに戻る時刻が存在するが、これらは各ユーザーの確率分布の積を取った場合には影響が小さいと考えられる。

6. まとめ

本研究では、各時刻で各ユーザーがデマに騙されている程度を表す確率分布の検討を二項分布を用いて行った。その結果、デマが拡散された直後に騙されている確率のピークが集中しやすい傾向がみられたが、特徴量とした語句・記号リストについては精度向上のための調整が必要と考えられる。また今後得られた分布をベイズ推定のための事前確率分布として適用することで、デマの拡散と思われる状態をいち早く検知できるよう検討を進める。

参考文献

- [1] 『「大阪地震でシマウマ脱走」のデマが拡散。警察が注意呼びかけ HUFFPOST』 https://www.huffingtonpost.jp/20-18/06/18/zebra-osaka_a_23461399/, (参照 2018-08-21).
- [2] 『「エアコン使うな、教委通達」講師の虚偽ツイートが拡散：朝日新聞デジタル』, <https://www.asahi.com/articles/ASL8B5QWQL8BPPTB002.html>, (参照 2018-08-21).
- [3] 池田 圭佑, 榊 剛史, 鳥海 不二夫, 栗原 聡, 『ロコミに着目した情報拡散モデルの提案及びデマ情報拡散抑制手法の検証』, 第 115 回 MPS 研究会, 2017.
- [4] 佐藤 和也, 佐野 幸恵, 山田 健太, 渡辺 隼史, 三浦 航, 高安 秀樹, 高安 美佐子, 『震災時におけるデマの拡散と制御に関するデータ解析とシミュレーション』, SICE 第 3 回社会システム部会研究会, 2013.
- [5] 浅野 孝夫, 『離散数学 —グラフ・束・デザイン・離散確率—』, サイエンス社, 2010.