

## SNSにおけるエコーチェンバー現象の分析と検出

古川友花<sup>†</sup>森博彦<sup>‡</sup>東京都市大学大学院総合理工学研究科<sup>†</sup>東京都市大学知能情報工学科<sup>‡</sup>

## 1. 研究背景

SNS(Social Networking Service)は現在、文字や画像、動画など多様な形態での情報の共有が行われ、共通の趣味をもつユーザとの交流等のツールとして支持されている。また、企業による広告を目的とした使用や、公共機関の災害等に関連した情報の提供も行われている。加えて、2013年の公職選挙法の一部を改正する法律の成立から、SNSを利用した選挙活動が行われていることや、「MeToo」運動などハッシュタグ(#)による連帯の表明によってオフラインの社会に影響を与えるような運動の広まりなど、議論の場としての役割を持ちつつあると考えられる。

SNSの特性から、ユーザは自身と同じ趣味や意見を持つユーザを見つけ、容易につながることが可能である。一方で、人間には、自身に肯定的な意見は強く捉え、否定的な意見は弱く捉えるBiased Assimilationや、自身の好む意見に接触し、好まない意見は避ける選択的接触といった心理的性質が存在する。これらの性質と結びつくことで、その性質を強化するという問題が起きており、その典型例として挙げられるのがエコーチェンバー現象(Echo Chambers)である。エコーチェンバー現象とは自身が持つまたは表明しているものと同様のコンテンツを消費する状況[1]を指す。意見の強化・増幅を生じさせ、攻撃的な意見の増加や誤情報の伝播[2]の原因であると言われており、その解決は重要な課題である。

## 2. 関連研究

Garimellaら[1]はTwitter上の政治議論に存在するエコーチェンバー現象とその構築について研究を行った。ユーザの投稿に含まれるニュースメディアのURLを基にユーザに極性をつけ、ユーザのコンテンツの生産と消費は二峰性の分布を示すことや、消費・生産極性は高い相関があることを発見し、政治議論におけるエコーチェンバー現象の存在を示唆した。極性に関係なく多様なコンテンツを共有するユーザはネットワ

ークにおける影響力やコミュニティとの繋がり  
の強さの点で、極性の高いユーザよりも劣ると  
分析している。

またVicarioら[2]は、Facebookにおける陰謀論と科学についての情報に関連したユーザのカスケードダイナミクスについて分析し、オンライン上の誤情報の拡散要因について調査を行った。両方の記事の拡散のパターンは異なっていたが、どちらのコンテンツも同質なクラスタの中で循環する傾向があることを示し、ユーザ間の同質性がコンテンツ拡散の駆動要因であることを明らかにした。

## 3. 目的

本研究では、Twitterを対象に、エコーチェンバー現象をユーザの特徴を使用し、検出することを目的とする。エコーチェンバーの状態にあるユーザを検出することによって、誤情報の拡散元を発見しやすくなることや、将来的なエコーチェンバー現象の改善手法の提案につながることを期待される。

## 4. ユーザの検出

## 4.1 分析とデータ

エコーチェンバー現象は対比のある分野で顕著であるため、関連研究と同様に、政治に関心のあるユーザを対象とした。2019年7月4日から7月19日の間に政治に関連したキーワードを含むツイートをTwitter上に5回以上投稿したユーザを政治に関心のあるユーザとしている。これらのユーザの最新ツイートから最大3200件をTwitter APIを用いて取得した。データを取得したユーザは497人で、計42,8552ツイートである。

これらのユーザのデータから、①ユーザのツイートの語彙数の時系列変化、②ユーザのツイートとフレンド・フォロワーとの文章の類似度、③フォロワー・フレンドの重複率、④ユーザのツイート間隔を算出し、これらを特徴量とした。①では、語彙に関する基礎的な指標であるTTR(Type Token Ratio)を用いた。TTRは文章の延べ語数をN、異なり語数をVとしたときに、 $TTR=V/N$ で求められる。ユーザの一週間のツイートのTTRと時系列の相関を算出し、それを使用した。エコーチェンバーの状態であれば、ユーザ

Analysis and detection of Echo Chambers on Social Networking Services

<sup>†</sup>Yuka Furukawa, Tokyo City University Graduate School

<sup>‡</sup>Hirohiko Mori, Tokyo City University

が同じ意見を持つユーザと交流していく過程で、ユーザの語彙は減少していくというような特徴があるのではないかと考えた。TTR を算出する際にはユーザのツイートを形態素解析ソフト MeCab[3] で形態素に分解し、助詞・助動詞は取り除いている。また、各ツイートの種類（オリジナルツイート、リプライ、リツイート (RT)）ごとに算出している。②では前述のユーザのフォロワー・フレンドからランダムに選んだ各 100 人のフォロワー・フレンドとのコサイン類似度を算出し、類似度が 0.53（ランダムなユーザ 5000 人 2500 組のコサイン類似度×4 回の平均値）を超えるユーザの割合を用いた。③では、ユーザがフォローしているユーザであるフレンド(A)と、ユーザをフォローしているユーザであるフォロワー(B)における重複した割合を使用しており、 $重複率 = \frac{A \cap B}{A + B - A \cap B}$  で算出される。フォロワーやフレンドとの意見の類似と併せ、ユーザが同様の意見を持つユーザたちと閉じた関係の中にいることの指標となると考えた。④では、ユーザのオリジナルツイート、リプライ、他のユーザが行うツイートの再投稿の形である RT の各ツイートの間隔の平均及び標準偏差を算出した。②や③に加えて、ユーザが発信したり、受け取る情報の速さや期間における量もユーザの意見に影響を与えらる。と考えられる。

#### 4.2 検出方法

エコーチェンバーの状態にあるユーザの検出には SVM (サポートベクターマシン) を使用した。前述の政治に関心のあるユーザに対して、エコーチェンバー状態にあるか否かについて付与したラベルと、各ユーザの①～④を特徴量として、学習データとしている。エコーチェンバーであるとラベルづけされたユーザは 308 人、エコーチェンバーでないとラベルづけされたユーザは 189 人である。

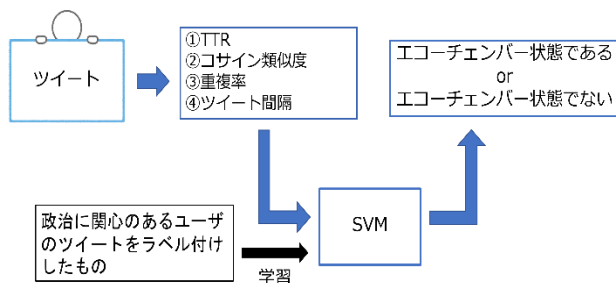


図 1. 検出の手順

#### 5. 結果

SVM による分類の結果、正解率=62.9%、適合率=64.7%、再現率=62.9%、F 値=63.5%となった。エ

コーチェンバーであると評価されたユーザのほうが正しく分類されやすかったが、全体として精度が高いとはいえない結果となった。

#### 6. 考察

正しく検出された内、(1)「エコーチェンバーである」とラベルづけされたユーザと、(2)「エコーチェンバーでない」とラベルづけされたユーザを比較すると、①ユーザのツイートの TTR の値の時系列変化では、(2)が低い相関の分布であった。しかし、④ツイート間隔では大きく差はなかったことから、ユーザの語彙の要素が、エコーチェンバー状態のユーザの検出に寄与していると考えられる。

また、ラベルづけを行う際に、人間がしている評価に、ユーザのツイートに対する感情的な面や話題性についての評価軸があり、それらを反映していないことによる評価のズレが生じている可能性が考えられる。

#### 7. 今後の課題

今後の課題としては、ユーザの発言が対立する話題に対してユーザの抱く感情の方向性や極性の指標についての検討があげられる。ユーザが対立する意見に対して攻撃的になることや、自分の意見への確信の度合等もエコーチェンバー状態の表出であると考えられる。しかし、新語表現の多い SNS では、従来のネガティブ・ポジティブ語を用いた処理だけではそれらの表現は難しいと考えられ、検討できていない。また、ユーザに対してのラベルを付与する際には、より多くの人に評価してもらい、一般性のあるものにする必要があるだろう。

#### 参考文献

- [1] Kiran Garimella, Gianmarco De Francisci Morales, Michael Mathioudakis: "Political Discourse on Social Media : "Echo Chambers, Gatekeepers, and the Price of Bipartisanship", In WWW 2018: The Web Aristides Gionis, Conference, pp. 913-922, 2018
- [2] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi: "The spreading of misinformation online", PNAS, vol. 113, no. 3, pp. 554-559, 2016
- [3] 形態素解析システム MeCab <https://taku910.github.io/mecab/>