

Twitter ユーザのパーソナリティ推定に基づく炎上現象の分析

○小川 祐樹 (立命館大学), 高 史明 (東京大学), 鳥海 不二夫 (東京大学)

キーワード: ネット炎上、Twitter、パーソナリティ、機械学習

1. はじめに

ネット上における炎上とよばれる集合現象は、個人・集団の意見や行動を過度に助長する危険性もあり大きな社会問題になりつつある。しかし、これらに関与する人々がどのようなパーソナリティ特性を持っているのかは、関与者の特性の測定が困難なことからほとんど検討されてこなかった。本研究では、ネット調査から得た Twitter ユーザのパーソナリティデータをもとに、機械学習によって炎上に関与する人々のパーソナリティ特性の推定を行う。そして、関与者の意見がどのような人達によって受容・拡散されていくかを明らかにする。

2. 関連研究

ネット炎上の関与者について、発言する人はインターネットユーザのごく一部であること(山口, 16)、保守系ユーザが関与している(鳥海, 17)という知見などが示されている。特に、保守性のユーザはネット上の差別的な発言を行うことが報告されており(高, 15)、保守性や社会支配志向が強いほど、古典的・現代的レイシズムを強めることが明らかになっている。この他、ネット利用量が多いほど攻撃性が高くなることや(高比良, 06)、ネット上で荒らし行為を行うユーザのサイコパシーが高いことなどが報告されている(Buckels, 14)。

差別的な言動や荒らしに関与するこれらの人々に対し、社会調査などで個人の負のパーソナリティ特性を測定することは、パーソナリティの特性からデータの信頼性の問題も生じる。このため、炎上現象の関与者がどのようなパーソナリティ特性を持っているかを定量的に示すことはほとんど検討されてこなかった。しかし、近年では情報学的アプローチとして、ネット利用者の行動データから機械学習で態度やパーソナリティを推定する研究が発展してきおり、ネット利用者の大規模なパーソナリティデータを収集することが可能になってきている(Golbeck, 11, Sumner, 12)。

本研究では、機械学習により Twitter ユーザのパーソナリティを行動データから推定する分類器を構築し、この分類器を用いて炎上関与者のパーソナリティ特性を推定することで、炎上関与者の特性を明らかにするというアプローチを取る。その際の仮説として、ネット炎上関与者は「党派

性(保守)が高い(または低い)」「攻撃性が高い」「ダークトライアド(サイコパシー)が高い」「社会支配志向」が高いかという点で分析を行う。

3. 分析方法

本研究では、機械学習によりパーソナリティを推定する分類器の構築と、炎上関与者のパーソナリティ推定の2つの手順をとる。研究アプローチの全体像を図1に示す。

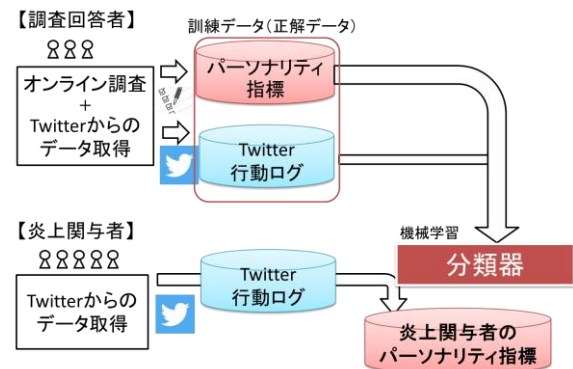


図1: 炎上関与者のパーソナリティ推定

3.1. Twitter ユーザのパーソナリティ推定

パーソナリティ指標については、ネット調査を利用して Twitter 利用者のアカウントと紐付けたデータの取得を行う。調査は2016年4月19日から4月22日に行い、スクリーニング条件「Twitter を利用している(週1回程度の閲覧・投稿)」「Twitter アカウントの公開設定」「Twitter アカウントのアプリ連携への許諾」「性別・年齢で均等割付(20歳未満, 20代, 30代, 40代, 50歳以上)」として調査を行い、1,592人のデータを得た。パーソナリティ指標を得るための質問項目は、ビッグファイヴ(小塩, 12)、ダークトライド(田村, 15)、攻撃性(安藤, 99; 高比良, 06)、社会支配志向(Pratto, 94; Ho, 12; Jost, 00; Kugler, 10)の質問項目によって測定した。

Twitter 行動データについては、被験者の Twitter アカウントから直近のツイート 3,200 件、フォロー関係のデータを Twitter API を用いて取得した。機械学習には RandomForest を用いて行う。形態素解析の辞書には Wikipedia の用語が入った NEologd、品詞として名詞・動詞・形容詞、特徴量とし

てtf・idfを用いた。また、4つのモデル (1. フォローのみ、2. ツイート (1gram)、3. ツイート (2gram)、4. フォロー+ツイート (2gram)) に対して機械学習の性能を MAE によって評価した (図2)。

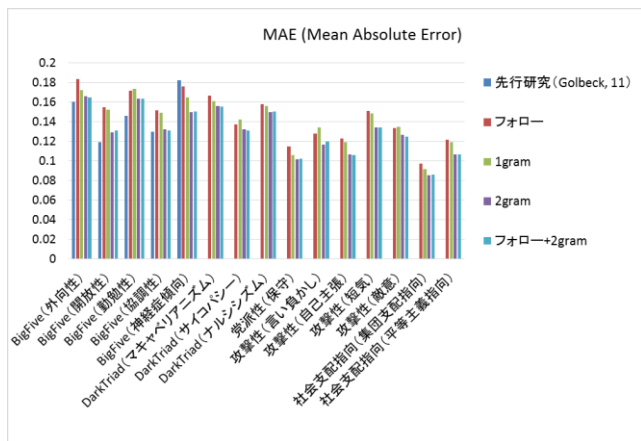


図2：機械学習の性能

図2の結果より、「フォロー+2gram」のモデルの性能が良く、ビッグファイブの推定性能は先行研究 (Golbeck, 11) とほぼ同程度であった。また、サイコパス、攻撃性、党派性 (保守)、社会支配志向ともにビッグファイブと同程度の性能で推定できることが分かった。

3.2. 炎上関与者のパーソナリティ推定

本稿では、炎上事例として「伊藤詩織」氏の炎上事例を取り上げる。これは、伊藤詩織氏が暴行を受けたとしてTBS元支局長の山口敬之氏を告発する会見を2017年5月29日に開き、Twitterを中心として複数の議論が起こった事例である。この事例において、Twitter上では伊藤氏の支援・応援・#MeTooの話題や、批判 (伊藤氏への批判 (疑問視、やりすぎ)、山口氏への批判)、メディア報道や捜査において政治的圧力があつたのではないかと複数の話題で多くのツイートがなされた事例である。

炎上の関与者の候補となるユーザとして、2017年5月29日から2018年1月31日の間において「伊藤詩織」を含むツイートをホットリンクのクチコミ係長を用いて取得した。総ツイート数は、81,807ツイート (うち、76,333がRT)、ユーザ数は28,961人であった。また、取得したツイートは複数の話題が混在するため、RT関係を用いたツイートの話題分類 (鳥海, 17)を用いて話題を抽出した。抽出されたトピックを表1に示す。トピックは「手記発売」「海外メディア報道」「支援・応援」「保守系ニュースまとめサイト」「捜査」の5つが抽出された。

表1：抽出されたトピック

トピック	ラベル	tf・idf上位用語 (トピックに含まれるRT)	トピック所属RT数
T1	手記発売	伊藤_詩織_詩織_さん_被害_告発_RT_NOSUKE_被害者_救済_shukan_shincho_捜査_司法_RT_tokyo_強姦_神話_さん_手記_手記_批判_二_次_次_加害_手記_Black_司法_システム	42
T2	海外メディア報道	伊藤_詩織_詩織_さん_RT_umekichkun_海外_メディア_よう_酷い_検証_超党派_イギリス_スウェーデン_メディア_報道_の_大きく_逮捕状_発布_追求_よう_山本太郎_件_なか_日本_石賀_雑誌_氏_事_NYT_伊藤	41
T3	支援・応援	伊藤_詩織_司法_問_詩織_さん_RT_kikko_kikko_no_no_blog_詩織_氏_自民党_公明党_救急外来_こと_大粒_涙_RT_iwakamiyasumi_安倍_総理_紹介_RT_特集_ページ	39
T4	保守系ニュースまとめサイト (アノニマスポスト)	RT_anonymous?_ネット_反応_伊藤_詩織_詩織_さん_柚木_護員_セハラ_件_しない_主張_YES_日本外国特派員協会_主張_の_詩織_弁護士_%_警察_記者_ソース_墓穴_www_これ_面白く	10
T5	捜査	伊藤_詩織_詩織_さん_捜査_員_事件_メディア_安倍晋三_首相_事件_日本_事件_安倍政権_喚問_無罪_RT_mt_格_元_よう_国_mt_mt_米_紙_元_刑事部長_何_ん	44

4. 結果

抽出された各トピックのツイートをRTしているユーザ群の推定パーソナリティ指標の平均を算出した。表2～表4は、訓練データユーザ群との差分を示したものである (「各トピックのツイートをRTしているユーザ群の推定パーソナリティ指標の平均」－「訓練データユーザ群のパーソナリティ指標平均」)。

表2は、ビッグファイブの結果である。この結果から、関与ユーザ全体 (T1～T5 ユーザ) において、外向性が高く、また動機性・協調性が低いことが分かる。つまり、関与ユーザはエネルギーであるが思いやりがなく衝動的なユーザであるといえる。

表3は、ダークトライドと攻撃性の結果である。この結果から、関与ユーザ全体においてダークトライド (サイコパス) が高く、また攻撃性が高い (特に、自己主張) が分かる。

表4は、党派性と社会支配志向の結果である。この結果から、関与ユーザの中でも、「保守系ニュースまとめサイト」のツイートのRTをするユーザの保守性が高く、社会支配志向 (集団支配志向主義) は高くないことが分かる。

表2：各トピックのツイートをRTするユーザのビッグファイブ

トピック	ラベル	BigFive (外向性)	BigFive (開放性)	BigFive (動機性)	BigFive (協調性)	BigFive (神経症傾向)
関与ユーザ	—	0.0272	+0.0029	-0.0286	-0.0665	+0.0114
T1ユーザ	手記発売	0.0300	+0.0093	-0.0268	-0.0667	+0.0133
T2ユーザ	海外メディア報道	0.0266	+0.0083	-0.0357	-0.0680	+0.0156
T3ユーザ	支援・応援	0.0337	+0.0065	-0.0268	-0.0707	+0.0082
T4ユーザ	保守系ニュースまとめサイト	0.0294	+0.0092	-0.0210	-0.0582	+0.0187
T5ユーザ	捜査	0.0119	+0.0052	-0.0376	-0.0680	+0.0069

表3：各トピックのツイートをRTするユーザの
ダークトライアドと攻撃性

トピック	ラベル	DarkTriad (マキヤベ リアニズム)	DarkTriad (ナルシシ ズム)	DarkTriad (サイコパ シー)	攻撃性 (言い負か し)	攻撃性 (自己主 張)	攻撃性 (短気)	攻撃性 (敵意)
関与ユーザ	—	-0.0589	0.0189	0.0457	0.0296	0.0735	0.0319	0.0372
T1ユーザ	手記発売	-0.0574	-0.0201	0.0468	0.0332	0.0969	0.0354	0.0384
T2ユーザ	海外メディア報道	-0.0609	0.0170	0.0516	0.0296	0.0786	0.0379	0.0397
T3ユーザ	支援・応援	-0.0705	0.0224	0.0460	0.0311	0.0628	0.0376	0.0391
T4ユーザ	保守系ニュースまとめサイト	-0.0583	-0.0224	0.0343	0.0379	0.0508	0.0221	0.0268
T5ユーザ	捜査	-0.0474	-0.0155	0.0499	0.0177	0.0695	0.0268	0.0419

表4：各トピックのツイートをRTするユーザの
党派性と社会支配志向

トピック	ラベル	党派性 (保守)	社会支配 指向(平 等主義指 向)	社会支配 指向(集 団支配指 向)
関与ユーザ	—	0.0427	0.0493	-0.0222
T1ユーザ	手記発売	0.0113	0.0597	-0.0339
T2ユーザ	海外メディア報道	0.0132	0.0558	-0.0311
T3ユーザ	支援・応援	0.0087	0.0573	-0.0305
T4ユーザ	保守系ニュースまとめサイト	0.1670	0.0236	0.0071
T5ユーザ	捜査	0.0111	0.0506	-0.0232

5. 考察

分析結果より、本論文で調査仮説としていた、炎上関与者は「保守的イデオロギーが高い（または低い）」「攻撃性が高い」「ダークトライアド（サイコパシー）が高い」については整合する結果が得られた。しかし、社会支配志向については、T4 ユーザのみが若干プラスとなる程度にとどまった。この理由として、今回の炎上事例特有の傾向の可能性が考えられる。今回の事例では、被害者が会見を開いて暴行被害を訴えるという事象に対し、平等主義のパーソナリティが高い人々が反応した可能性があり、炎上で扱われる話題によって変わることが考えられる。今回は一つの事例に対する分析のみであったが、多数の炎上事例を分析し、共通するパーソナリティ指標、また話題特有のパーソナリティ指標の可能性について検討していく必要がある。

6. まとめ

本研究では、機械学習により Twitter ユーザのパーソナリティを行動データから推定するモデルを構築し、炎上関与者のパーソナリティ特性を明らかにした。その結果、炎上関与者はダークトライアド（サイコパシー）が高く、攻撃性が高く、保守系ニュースまとめサイトを RT するユーザ群において保守性が高いことが明らかになった。今後の課題として、様々な炎上事例の分析し、一般化や個別事象の特性について検討していく。

【参考文献】

- [安藤 99] 安藤明人・曾我祥子・山崎勝之・島井哲志・嶋田洋徳・宇津木成介・大芦治・坂井明子 (1999). 日本版 Buss-Perry 攻撃性質問紙 (BAQ) の作成と妥当性、信頼性の検討, 心理学研究, Vol.70, pp.384-392
- [Buckels 14] Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences*, 67, 97-102.
- [Golbeck 11] Golbeck, J., Robles, C., Edmondson, M. and Turner, K. (2011). Predicting Personality from Twitter, IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom) : 149-156.
- [Ho 12] Ho, A. K., Sidanius, J., Pratto, F., Levin, S., Tomsen, L., Kteily, N., & Sheehy-Skefington, J. (2012). Social dominance orientation: Revisiting the structure and function of a variable predicting social and political attitudes. *Personality and Social Psychology Bulletin*, 38, pp.583-606.
- [Jost 00] Jost, J. T. & Tompson, E. P. (2000). Group-based dominance and opposition to equality as independent predictors of self-esteem, ethnocentrism, and social policy attitudes among African Americans and European Americans. *Journal of Experimental Social Psychology*, Vol.36, pp.209-232
- [Kugler 10] Kugler, M. B., Cooper, J., & Nosek, B. A. (2010). Group-based dominance and support for equality as consequences of different psychological motives. *Social Justice Research*, Vol.23, pp.117-155.
- [小塩 12] 小塩真司, 阿部晋吾, カトローニ ピノ (2012). 日本語版 Ten Item Personality Inventory (TIPI-J)作成の試み パーソナリティ研究, 21, 40-52.
- [Pratto 94] Pratto, Felicia, James Sidanius, Lisa M. Stallworth, and Bertram F. Malle (1994). Social Dominance Orientation: A Personality Variable Predicting Social and Political Attitudes, *Journal of Personality and Social Psychology*, 67, 4, pp.741-763
- [Sumner 12] Sumner, C., Byers, A., Boochever, R. & Park, G. J (2012). Predicting Dark Triad Personality Traits from Twitter Usage and a Linguistic Analysis of Tweets, ICMLA '12 Proc. Of 2012 11th International Conference on Machine Learning and Applications, 02, 386-393.
- [高 15] 高史明. (2015). レイシズムを解剖する：在日コリアンへの偏見とインターネット, 勁草書房
- [田村 15] 田村紋女, 小塩真司, 田中圭介, 増井啓太, ジョナソン ピーター カール (2015). 日本語版 Dark Triad Dirty Dozen (DTDD-J) 作成の試み パーソナリティ研究, 24, 26-37.
- [高比良 06] 高比良美詠子・安藤玲子・坂元章 (2006). 縦断調査による因果関係の推定 -インターネット使用と攻撃性の関係, パーソナリティ研究, Vol.15, No.1, pp.87-102.
- [鳥海 17] 鳥海不二夫, 榎剛史. (2017). バースト現象におけるトピック分析, 情報処理学会論文誌, 58, 6, 1287-1299.
- [山口 16] ネット炎上の研究, 田中辰雄, 山口真一. (2016). ネット炎上の研究, 勁草書房.