The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

2017年衆議院選挙における政党公式アカウントフォロワーの分析

Analysis of Political Party Accounts' Followers in Japan Electoin 2017

鳥海不二夫 *1　吉田光男 *2
Fujio Toriumi　Mitsuo Yoshida

*1東京大学　*2豊橋技術科学大学
The University of Tokyo　Toyohashi University of Technology

In the election campaign, modern political parties utilize social media to advertise their policies and candidates and to communicate electorates. In the latest election, the 48th general election for the Lower House in Japan 2017, social media, especially Twitter, was actively used. In this paper, we perform a detailed analysis of political tweets of users who retweeted offical political party accounts during the election. Our aim is to obtain how political party followers differs from each other. The results indicate that the largest two political party accounts' followers are quiet similar from view point of political tweets.

1. 緒言

ネット選挙が解禁されて依頼，日本においても選挙時に Twitter や Facebook といったソーシャルメディアを利用することは，支持者への情報発信や，新たな支援者の獲得に欠かせないものとなっている．一方で，ネット選挙の影響は限定的であるという報告も存在し [総務 14]，日本においてどのようなネット選挙の戦略が有効であるかはいまだ明らかになっていない．

ネット選挙では，ソーシャルメディアにおける拡散の強さに期待が集まっている．特に，ソーシャルメディアでは従来メディアとは異なりプッシュ型の情報拡散を行うことが可能であるため，元来当該情報に興味がないユーザに対しても他のユーザのリツイートなどの拡散行動によって情報を伝えることが可能である．そのため，各政党の公式アカウントの発言が他のユーザの協力によってソーシャルメディア全体に拡散し，より多くの人々に情報が伝わることが期待される

そこで，本研究では 2017 年衆議院選挙時の各政党公式アカウントのフォロワーについて，その特徴を明らかにすることを目指し分析を行う．本稿では，公式アカウントのツイートをリツイートするユーザが，政治的関心が高いユーザか一般的なユーザかによって情報拡散の質が異なるであろうことを考慮し，各政党のツイートがどのようなユーザによってリツイートされていたのかを分析する．

もし，日常的には政治的発言が少ないユーザによってもリツイートされている政党公式アカウントのツイートがあれば，それは一般のユーザにまでその情報が波及した可能性が高いと考えられるため，非支持層へのアピールとなるだろう．一方，日常的に政治的発言が多いユーザは，もともと当該政党の支持者である可能性が高いため，そのようなユーザによるリツイートが多いツイートは支持者へのアピールに寄与していると考えられる．

このような仮説の元，各政党の公式アカウントのツイートをリツイートしたユーザを分析することで，ツイッター戦略の違いを明らかにすることを目指す．

連絡先: 鳥海不二夫，東京大学，東京都文京区本郷 7-3-1，tori@sys.t.u-tokyo.ac.jp

2. 関連研究

Tumasjan らは Twitter が政治のためのディスカッションフォーラムとして機能していると報告している [Tumasjan 10]．一方で，ユーザが政治情報へアクセスすることについて分断が存在することが示されている [Adamic 05, Hayat 17, Hyun 14, Iyengar 09]．このような調査は主に，ソーシャルメディア上での政治に関するニュース記事について調査しており，政党のツイートなどを取り扱っているわけではない．

一般的に，フォロワー数はユーザの注目度として用いられるが，フォロワーを詐欺的に獲得するケースも存在し [Stringhini 13]，フォロワー数が必ずしも情報拡散に影響しているとは言えない．実際，政治的なコミュニケーションにおいて，情報のハブはフォロワー数だけでは決まらないことが明らかになっている [Bakshy 11]．

一方で，ユーザのツイート特性に基づいて各政党のツイート戦略について分析した研究はほとんどない．

3. データセット

3.1 収集データ

本研究で利用するために収集したデータの収集期間とその量を表 1 に示す．

主要 6 政党（自民党，立憲民主党，希望の党，公明党，共産党，日本維新の会）の公式アカウントについて，そのツイートをリツイートしたユーザを収集した．さらに，公式アカウントのツイートをリツイートしたユーザについて，過去にさかのぼってのツイートを取得 *1 した．

3.2 公式アカウントの基本データ

各政党の公式アカウントの基本情報を表 2 に示す．

ここで，Tweets は公式アカウントが行ったツイート数，Retweets は公式アカウントのツイートがリツイートされた総数，Follower はアカウントをフォローしているユーザ数，Friends は公式アカウントがフォローしているユーザ数である．

これより，多くの政党アカウントは Follower は多いが，自身はほとんど Follow を行っていないことがわかる．また，ツイート数には最大で 3 倍程度の差があり，ツイッターを積極的

*1 Twitter の API 制限により最大 3200 ツイートまで収集している

表 1: データセット

	内容	収集期間	総数
1	Official Accounts	-	6
2	OA's Tweet	2017/9/27-10/23	2790
3	Retweeted Users	2017/9/27-10/23	84043
4	Users' Tweet	2007/04/10-2017/11/14	197,267,467

に利用していた政党と，あまり積極的には利用していなかった政党があったといえる．

また，リツイートされた総数をみると立憲民主党が極めて多く，自由民主党がそれに続いており，この二つの政党の情報が多く共有されていたことがわかる．

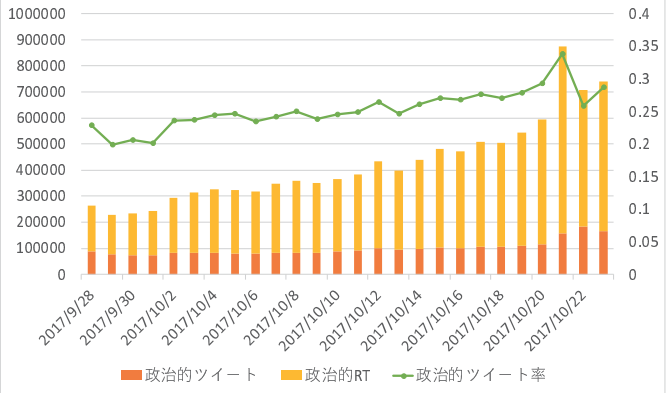

図 1: 選挙期間内の政治的ツイート

4. 政党的発言の判別

4.1 政治的ツイートの学習

各ユーザがどの程度政治的発言を行ったり，政治的発言をリツイートしているかを調べるため，各ツイートが政治的発言かどうかを機械学習を用いて判別する．

まず，人手でランダムに抽出した 9500 ツイートを，政治的ツイートか一般的なツイートかに分類した．次に，分類したツイート群を教師データとし，SVM[Vapnik 63]，ロジスティック回帰 [Cox 58]，CNN[Kim 14] を用いて学習を行った．

SVM, ロジスティック回帰では TF-IDF によって正規化した BugOfWords に変換し学習し，CNN では Tweet 内の単語を FastText により求めた単語分散表現に変換し，畳み込みニューラルネットワークを用いて学習を行った．

4.2 政治的内容を含むツイートの抽出

3 種類の手法について，全体の 70%を学習データ，30%をテストデータとして用い学習を行った．その結果得られた，テストデータに対する精度を表 3 に示す．

これより，CNN を用いた手法によって最も高い精度で政治的発言が抽出できることが確認できたた．

次に，収集した全ツイートに対して政治的発言かどうかの判別を行った．なお，ここで収集した全ツイートとは，公式アカウントのツイートをリツイートしたユーザのツイート群である．その結果，11,635,182 ツイートが政治的ツイートであると判別された．

また，選挙期間中 (9 月 28〜10 月 23 日) において，政治的ツイートおよび政治的ツイートのリツイート，そして全体に占める政治的ツイート，リツイートの割合を図 1 に示す．これより，衆議院解散日から徐々に政治的な発言は増加していき，選挙前日にピークを迎えていることがわかる．ピークが選挙当日ではないのは，選挙当日はネット選挙を行うことができないためである．以上の結果からも，判別結果は妥当であると考えられる．

5. 公式アカウントをリツイートしたユーザの分析

5.1 リツイートユーザの重複

各政党の公式アカウントをリツイートしたユーザは，当該政党の政策になんらかの興味を持っているユーザであるといえる．そのようなユーザの特徴を分析することで，各政党アカウントの情報がどのようなユーザによって拡散されたかを明らかにする．以下，各政党のツイートをリツイートしたユーザを，リツイートユーザと呼ぶ．

まず，どの程度のユーザが複数政党のリツイートユーザとなっていたかを確認するため，二つの政党を同時にリツイートしたユーザの割合を確認した．その結果を表 4 に示す．この表では，左辺の政党リツイートユーザのうち，何パーセントが上辺の政党のリツイートユーザであるかを示している．

この結果から，多くのユーザが単一の政党のツイートしかリツイートしていないことがわかる．例外的に他の政党をリツイートしている割合が高いのは，

- 希望の党，共産党→立憲民主党
- 公明党，日本維新の会→自由民主党

である．立憲民主党の成立過程を考慮すれば，希望の党のリツイートユーザが立憲民主党のツイートをリツイートすることは妥当であろう．また，自由民主党と公明党も連立政権を組んでいることから，双方をリツイートする行動も妥当であると考えられる．

一方で，共産党のユーザの 70% が立憲民主党のツイートをリツイートしている点は興味深い．リツイートユーザの人数比が大きく異なることを考慮すると，この結果は共産党のリツイートユーザの多くが立憲民主党のツイートもリツイートしていたということになる．

以上，リツイートユーザの重複関係を分析することで，政党間のユーザ関係が明らかになる可能性が示唆された．

5.2 政治的発言を用いたユーザのクラスタリング

各政党のリツイートユーザについて，政治的発言の含有率に基づくクラスタリングを行った．これによって，各政党のリツイートユーザの特徴を分析する．

まず，ここでは各ユーザの選挙期間中のツイートについて，

- ツイート数 (TW)
- リツイート数 (RT)
- 政治的ツイート数 (PT)
- 政治的ツイートのリツイート数 (PRT)

から，

- リツイート率 (RT/TW)
- ツイート中政治的ツイート率 (PT/(TW-RT))
- 政治的リツイート率 (PRT/RT)

表 2: 公式アカウント情報・ツイートは 9/28-10/23．Follower,Friend 数は 11 月 10 日に収集

政党	アカウント名	Tweets	Retweets	Followers	Friends
自由民主党	jimin_koho	280	110685	134595	322
立憲民主党	cdp2017	904	506432	191011	77
希望の党	kibounotou	410	21559	13529	152
公明党	komei_koho	289	31072	76743	1259
日本共産党	jcp_cc	347	48203	42508	253
日本維新の会	osaka_ishin	281	13163	15999	177

表 3: 学習機の精度評価

学習機	精度
SVM	0.671
ロジスティック回帰	0.727
CNN	0.839

を求めユーザの特徴量とし，K-Means($k = 7$) によってユーザをクラスタリングを行った．

得られたクラスタを表 5 に示す．この中で特徴的なクラスタである C1,C2,C4,C6 について述べる．

C1 はリツイートが少なく，政治的発言も少ないが，リツイート内には政治的内容が 83.5% 含まれている．C2 はリツイートが少なく，政治的発言も少ないクラスタである．C4 はリツイートが多く，政治的なツイートリツイートともに少ないクラスタである．C6 はリツイートが多い一方で，自身のツイートの中に政治的発言が特に多いクラスタである．

5.3 各政党におけるリツイートユーザのクラスタ

各政党のリツイートユーザが，どのクラスタに所属しているのかを分析した．その結果を図 2 に示す．ここでは，政党ごとに各クラスタのユーザがどの程度の割合で存在しているかを示している．

これを見ると多くの政党で類似した傾向を持っていることがわかる．そこで，政党ごとの特徴を明確にするために，各政党のリツイートユーザ割合の相関を求めた．もし，相関が高ければ類似したユーザ群をリツイートユーザに持っていることになり，相関が低ければリツイートユーザの性質が異なることを意味する．

その結果が表 6 である．これより，まず自由民主党と立憲民主党は相関係数が 0.816 と極めて高いことがわかる．ここから，これら二つの政党のリツイートユーザは類似しているということを意味し，どちらかが突出した特徴を持っているとはいえないことが明らかとなった．

また，自由民主党と日本維新の会が類似しており，立憲民主党と公明党が類似した性質を持つ．表 4 より，これらのペアはいずれもリツイートユーザの重複は少ないことから，全く異なるユーザで構成されているにも関わらず，その性質は類似しているという点は興味深い．

一方，日本共産党のリツイートユーザは日本維新の会以外の政党のリツイートユーザと大きく異なる性質を持っていることがわかった．

ユーザ構成の違いが政党のツイッター戦略とどのように関連するかは今後さらなる分析が必要であるが，フォロワーの多い

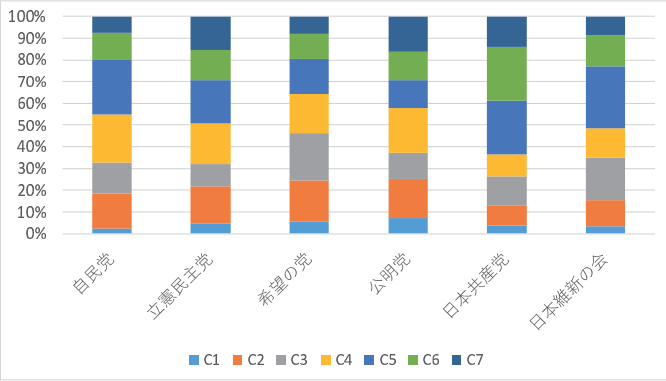

図 2: リツイートユーザの所属クラスタ

二政党の分布が類似している点は興味深い．立憲民主党のフォロワーの増加率は異常であるという言説も見受けられたが，この結果を見る限り，少なくともリツイートを行ったユーザについては，特に異常な値を示しているものではないといえよう．

6. 結言

本研究では，2017 年に行われた衆議院選挙の政党公式アカウントをリツイートしたユーザについて，分析を行った．リツイートユーザのツイートを政治的ツイートかそうではないかに分類し，その割合に基づいて，ユーザのクラスタリングを行った．さらに，各政党のリツイートユーザに各クラスタに所属するユーザがどの程度の割合含まれるかを分析し，その特徴を明らかにした．

その結果，自由民主党と立憲民主党のリツイートユーザの性質は類似していること，日本共産党のリツイートユーザは他の政党とは異なる性質を持つことなどが明らかとなった．

今後，リツイートユーザをさらに分析することで各政党に適したツイート戦略が明らかになると期待される．例えば，政治的ツイートのリツイート率が高い政党であれば，政治的ツイートを数多く行うことで，効率的に情報を拡散できると期待できる．ただし，拡散率はリツイートユーザのフォロワー率などにも影響されるため，フォローフォロワー関係などネットワーク構造についても考慮した分析が必要になると考えられる．

参考文献

[Adamic 05] Adamic, L. A. and Glance, N.: The political blogosphere and the 2004 U.S. election: divided they blog, in *LinkKDD '05*, pp. 36–43 (2005)

表 4: 公式アカウントのリツイート重複率

	自民党	立憲民主党	希望の党	公明党	共産党	維新の会
自民党		0.101	0.032	0.104	0.035	0.033
立憲民主党	0.043		0.021	0.022	0.148	0.005
希望の党	0.187	0.285		0.093	0.110	0.056
公明党	0.222	0.110	0.034		0.042	0.023
共産党	0.070	0.708	0.038	0.040		0.009
維新の会	0.380	0.131	0.111	0.126	0.050	

表 5: 政治的発言に基づくユーザクラスタ

	RT/TW	PT/(TW-RT)	PRT/RT	ユーザ数
C1	0.304	0.402	0.835	1991
C2	0.332	0.110	0.167	7662
C3	0.421	0.576	0.391	4690
C4	0.825	0.094	0.159	8772
C5	0.844	0.425	0.391	9081
C6	0.865	0.809	0.498	6204
C7	0.951	0.032	0.564	6617

表 6: 政治的発言に基づくユーザクラスタの類似性

	自民党	立憲民主党	希望の党	公明党	日本共産党	日本維新の会
自民党		0.816	0.731	0.539	0.470	0.827
立憲民主党	0.816		0.454	0.801	0.471	0.557
希望の党	0.731	0.454		0.455	0.137	0.624
公明党	0.539	0.801	0.455		0.004	0.097
日本共産党	0.470	0.471	0.137	0.004		0.731
日本維新の会	0.827	0.557	0.624	0.097	0.731	

[Bakshy 11] Bakshy, E., Hofman, J. M., Mason, W. A., and Watts, D. J.: Everyone's an Influencer: Quantifying Influence on Twitter, in *WSDM '11*, pp. 65–74 (2011)

[Cox 58] Cox, D. R.: The regression analysis of binary sequences, *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 215–242 (1958)

[Hayat 17] Hayat, T. and Samuel-Azran, T.: "You too, Second Screeners?" Second Screeners' Echo Chambers During the 2016 U.S. Elections Primaries, *Journal of Broadcasting & Electronic Media*, Vol. 61, No. 2, pp. 291–308 (2017)

[Hyun 14] Hyun, K. D. and Moon, S. J.: News Media's Role in the Issue-Voting Process: News Attention, Issue Proximity, and Vote Choice, *Journalism & Mass Communication Quarterly*, Vol. 91, No. 4, pp. 687–705 (2014)

[Iyengar 09] Iyengar, S. and Hahn, K. S.: Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use, *Journal of Communication*, Vol. 59, No. 1, pp. 19–39 (2009)

[Kim 14] Kim, Y.: Convolutional neural networks for sentence classification, *arXiv preprint arXiv:1408.5882* (2014)

[Stringhini 13] Stringhini, G., Wang, G., Egele, M., Kruegel, C., Vigna, G., Zheng, H., and Zhao, B. Y.: Follow the Green: Growth and Dynamics in Twitter Follower Markets, in *IMC '13*, pp. 163–176 (2013)

[Tumasjan 10] Tumasjan, A., Sprenger, T., Sandner, P., and Welpe, I.: Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment, in *ICWSM '10*, pp. 178–185 (2010)

[Vapnik 63] Vapnik, V.: Pattern recognition using generalized portrait method, *Automation and remote control*, Vol. 24, pp. 774–780 (1963)

[総務 14] 総務省：インターネット選挙運動解禁に関する調査報告書 (2014)