

Will Sanders Supporters Jump Ship for Trump? Fine-grained Analysis of Twitter Followers

Yu Wang
Political Science
University of Rochester
Rochester, NY, 14627
yu.wang@rochester.edu

Yang Feng
Computer Science
University of Rochester
Rochester, NY, 14627
yfeng23@cs.rochester.edu

Xiyang Zhang
Psychology
Beijing Normal University
Beijing, 100875
zxy2013@mail.bnu.edu.cn

Richard Niemi
Political Science
University of Rochester
Rochester, NY, 14627
niemi@rochester.edu

Jiebo Luo
Computer Science
University of Rochester
Rochester, NY, 14627
jluo@cs.rochester.edu

ABSTRACT

In this paper, we study the likelihood of Bernie Sanders supporters voting for Donald Trump instead of Hillary Clinton. Building from a unique time-series dataset of the three candidates' Twitter followers, which we make public here, we first study the proportion of Sanders followers who simultaneously follow Trump (but not Clinton) and how this evolves over time. Then we train a convolutional neural network to classify the gender of Sanders followers, and study whether men are more likely to jump ship for Trump than women. Our study shows that between March and May an increasing proportion of Sanders followers are following Trump (but not Clinton). The proportion of Sanders followers who follow Clinton but not Trump has actually decreased. Equally important, our study suggests that the jumping ship behavior will be affected by gender and that men are more likely to switch to Trump than women.

CCS Concepts

•Human-centered computing → Social engineering (social sciences); Social media;

Keywords

Presidential Election; Bernie Sanders; Donald Trump; Hillary Clinton; Gender;

1. INTRODUCTION

Even as Hillary Clinton moves closer and closer to clinch the Democratic nomination, some Sanders supporters insist that they will not vote for her in the general election. This leaves a golden opportunity to Donald Trump, the pre-

sumptive Republican nominee who has been sharing similar campaign messages with Sanders on trade and campaign finance. Trump, on the other hand, also makes it clear that he will target Sanders supporters.¹

This new dynamic quickly became hotly debated and the looming question is “Can Trump win over Sanders supporters?”² A number of recent polls, including ABC/Washington Post, CBS/NYT and YouGov, do suggest that some Sanders supporters could end up voting for Trump and that this is particularly so for his male supporters.³ In this paper, we investigate this dynamic in Twitter.

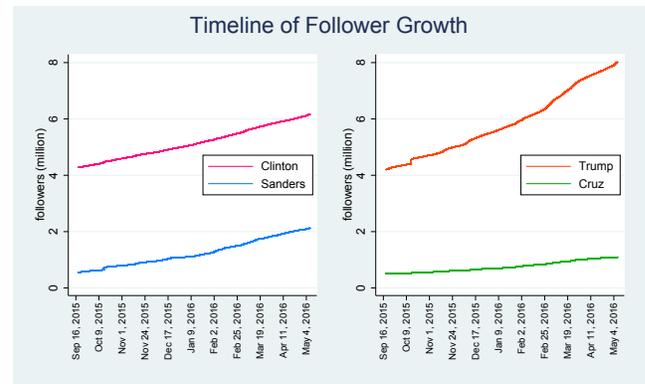


Figure 1: Number of Followers for Hillary Clinton, Bernie Sanders, Donald Trump and Ted Cruz.

Building from a unique dataset of the three candidates' Twitter followers (Figure 1), we first examine the proportion of Sanders followers who simultaneously follow Trump or Clinton. Then we train a convolutional neural network to classify the followers' gender and study how men in particular are responding to Trump's overture. Our study sug-

¹Newsweek, <http://www.newsweek.com/can-trump-win-over-bernie-sanders-supporters-459218>.

²New York Times, <http://www.nytimes.com/2016/04/29/us/politics/hillary-clinton-donald-trump-women.html>.

³New York Times, <http://www.nytimes.com/2016/05/25/upshot/explaining-hillary-clintons-lost-ground-in-the-polls.html>.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Under Review for CIKM 2016, Indiana USA

© 2018 Copyright held by the owner/author(s).

ACM ISBN 123-4567-24-567/08/06...\$15.00

DOI: 10.475/123_4

gests that in Twitter there also exists a shift towards Trump among Sanders followers, and that male Sanders followers are responding more positively to Trump than females.

2. RELATED LITERATURE

Our work builds on previous literature in electoral studies, data mining, and computer vision.

In electoral studies, researchers have argued that gender constitutes an important factor in voting behavior. One common observation is that women tend to vote for women, which is usually referred to as gender affinity effect [4, 1]. In the 2016 presidential election, Hillary Clinton also portrays herself as a champion “fighting for women’s healthcare and paid family leave and equal pay.” Our work will test the strength of this gender affinity effect when Sanders supporters decide whether to jump ship for Trump or not.

In data mining, there is a burgeoning literature on using social media data to analyze and predict elections. In particular, several studies have explored ways to infer users’ preferences. According to [7], tweets with sentiment can potentially serve as votes and substitute traditional polling. [12] exploits the variations in the number of ‘likes’ of the tweets to infer Trump followers’ topic preferences. [6] uses candidates’ ‘likes’ in Facebook to quantify a campaign’s success in engaging the public. [11] uses follower growth on public debate dates to measure candidates’ debate performance. Our work also pays close attention to the number of followers, but we go further by investigating the composition of these followers.

Our work also ties in with current computer vision research. In this dimension, our work is related to gender classification using facial features. [5] uses a five-layer network to classify both age and gender. [3] introduces a dataset of frontal-facing American high school yearbook photos and uses the extracted facial features to study historical trends in the U.S. [2] provides a comprehensive survey of race classification based on facial features. [9] uses user profile images to study and compare the social demographics of Trump followers and Clinton followers. [10] focuses specifically on the unfollowers of Donald Trump and Hillary Clinton and reports that women are more likely to unfollow both candidates. [8] studies the ‘woman card’ exchange between Trump and Clinton, and finds that the ‘woman card’ exchange has made women more likely to follow Clinton and less likely to unfollow her.

3. DATA AND METHODOLOGY

In this section, we describe our dataset *US2016*, the pre-processing procedures and our CNN model. One key variable is *number of followers*. This variable is available for all three candidates and covers the entire period from Sept. 18, 2015 to May 7, 2016. Compared with the candidates who have dropped out of the race, the three remaining presidential candidates also have the most Twitter followers (Figure 1). This variable is updated every 10 minutes.

Besides the number of followers, our dataset *US2016* also contains the detailed follower IDs for Trump, Clinton and Sanders on specific dates, including March 24th, April 17th and May 10th. This information enables us to track the evolution of the election dynamics. We report the summary statistics in Table 1.

Using follower information, we first study among Sanders

Table 1: Number of Followers

| | March 24 | April 17 | May 10 |
|-----------------|-----------|-----------|-----------|
| Bernie Sanders | 1,777,861 | 1,977,982 | 2,134,917 |
| Hillary Clinton | 5,755,618 | 5,905,124 | 6,176,731 |
| Donald Trump | 7,075,507 | 7,604,915 | 8,020,568 |

followers who are following Trump but not Clinton and who are following Clinton but not Trump. We think it is reasonable to assume that if Sanders drops out of the race, Sanders followers who are following Trump but not Clinton will support Trump, and that those who are following Clinton but not Trump will support Clinton. To match the candidates’ millions of followers, we first sort their IDs and then use binary search. The entire matching process can be done within a few minutes.⁴

Furthermore, we collect the profile images based on follower IDs. Our goal is to infer an individual’s gender based on the profile image and to test the hypothesis that individuals who follow both Sanders and Trump are more likely to be male than an average Sanders follower.

To process the profile images, we first use OpenCV to identify faces, as the majority of profile images only contain a face.⁵ We discard images that do not contain a face and the ones in which OpenCV is not able to detect a face. When multiple faces are available, we choose the largest one. Out of all facial images thus obtained, we select only the large ones. Here we set the threshold to 18kb. This ensures high image quality and also helps remove empty faces. Lastly we resize those images to (28, 28, 3). Eventually, we get 40,088 images sampled from all Sanders followers and 34,921 images from Sanders followers who also follow Trump (Table 2).

Table 2: Number of Profile Images in *US2016*

| | All Sanders | Sanders & Trump |
|------------------|-------------|-----------------|
| Number of Images | 40,088 | 34,921 |

To classify profile images, we train a convolutional neural network using 42,554 weakly labeled images, with a gender ratio of 1:1. These images come from Trump’s and Clinton’s current followers. And we infer their labels using the followers’ names. For example, David, John, Luke and Michael are male names, and Caroline, Elizabeth, Emily, Isabella and Maria are female names.⁶ For validation, we use a manually labeled data set of 1,965 profile images for gender classification. The validation images come from Twitter as well so that we can avoid the cross-domain problem. Moreover, they do not intersect with the training samples as they come exclusively from individuals who unfollowed Hillary Clinton before March 2016.

Table 3: Summary Statistics of CNN Performance

| Architecture | Precision | Recall | F1 | Accuracy |
|--------------|-----------|--------|-------|----------|
| 2CONV-1FC | 91.36 | 90.05 | 90.70 | 90.18 |

⁴Codes and data used in this paper are available on the first author’s website.

⁵<http://opencv.org/>.

⁶The full list of label names together with the validation data set and the trained model, is available at the first author’s website.

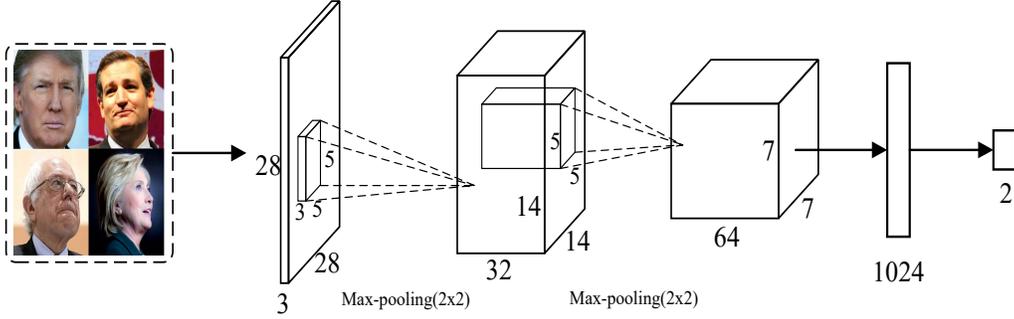


Figure 2: The CNN model consists of 2 convolution layers, 2 max-pool layers and a fully connected layer.

The architecture of our convolutional neural network is illustrated in Figure 2, and the performance of the model is reported in Table 3.

4. MAIN RESULTS

In this section, we answer the question of whether Sanders followers are jumping ship for Trump. Specifically, we examine whether an increasing proportion of Sanders followers are now following Trump and, if so, whether this phenomenon is particularly significant for men.

4.1

In this subsection, we analyze the composition of Sanders followers between March and May. We divide Sanders followers into four groups: (1) only follow Trump and Sanders, (2) only follow Clinton and Sanders, (3) follow Trump, Clinton and Sanders, (4) only follow Sanders. We assume that followers in Group 1 are the most likely to switch to Trump and followers in Group 2 are the least likely. We report our results in Figures 3, 4, 5.

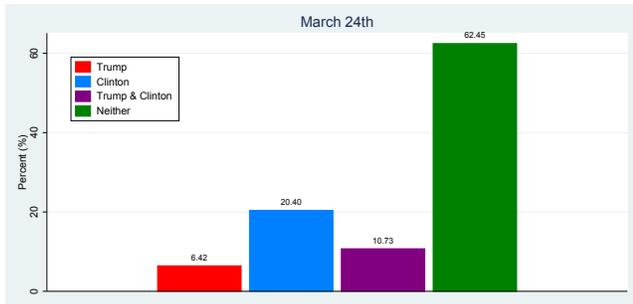


Figure 3: Composition of Sanders followers, March.

The results indicate a decrease of Clinton followers and an increase of Trump followers among Sanders followers between March and May. In Table 5, we use score test to show that the increase of Trump’s presence and the drop of Clinton’s presence are statistically significant.⁷

Meanwhile, we also observe that individuals who follow only Sanders, marked by green, make up a smaller share

⁷The formula for the score test statistic is: $z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)}}$, where $\hat{p}_1 = \frac{x}{n_1}$, $\hat{p}_2 = \frac{y}{n_2}$, $p = \frac{x+y}{n_1+n_2}$. With large n_1 and n_2 , z is approximately standard normal.

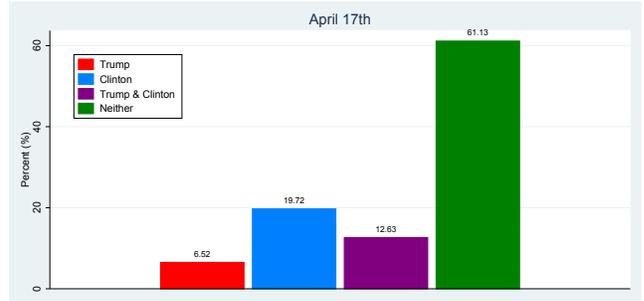


Figure 4: Composition of Sanders followers, April.

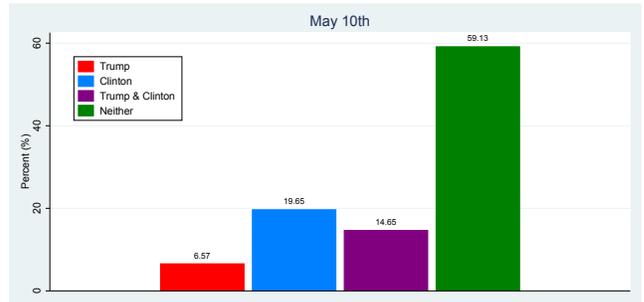


Figure 5: Composition of Sanders followers, May.

in May than in March and that the share of individuals who follow all the three candidates has increased. Using score test, we are also able to show that these changes are statistically significant.

Table 4: The Composition of Sanders Followers

| Null Hypothesis | Clinton & Sanders | | Trump & Sanders | |
|---------------------------|-------------------|---------|-----------------|---------|
| | z statistic | p value | z statistic | p value |
| $\Pr_{March} = \Pr_{May}$ | -18.47 | 0.00 | 5.99 | 0.00 |

4.2

In this subsection, we study whether men are more likely to jump ship for Trump than women. Our investigation is motivated by poll findings that show white male supporters

of Bernie Sanders are the most likely to switch to Trump.⁸

We use data collected on May 10th, when Sanders has 2,134,917 followers, of which 140,185 simultaneously follow Trump but not Clinton. Using the neural network reported in Section 3, we classify the gender of these followers. We report the results in Figure 6.

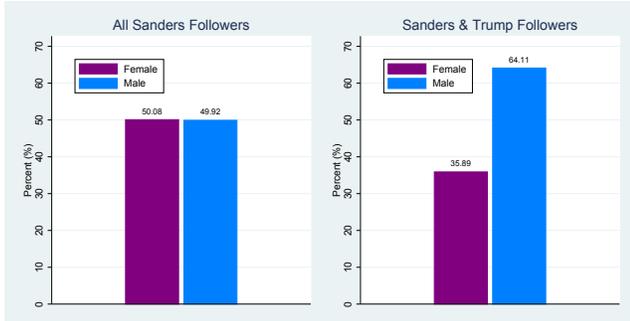


Figure 6: Sanders followers who are likely to jump ship for Trump are disproportionately male.

We find that of all Sanders followers 49.92% are male, but for those who follow both Sanders and Trump (and not Clinton) the percentage is as high as 64.11%. Using score test (Table 5), we show that among the Sanders followers, those who simultaneously follow Trump but not Clinton are more likely to be male than an average Sanders follower. Our results are consistent with the poll results and lend further support to previous studies that demonstrate the gender effect.

Table 5: Score Test on Gender Composition

| Null Hypothesis | Men | |
|--|-------------|---------|
| | z statistic | p value |
| $\Pr_{Sanders} = \Pr_{Sanders \& Trump}$ | 39.10 | 0.00 |

5. CONCLUSIONS

As Hillary Clinton moves closer and closer to clinch the Democratic nomination, the question of whether Sanders supporters would jump ship for Trump becomes decisive. A number of polls suggest that Sanders supporters could end up voting for Trump and that this is particularly so for his male supporters. In this paper, we explored this new dynamic in social media. Building from a unique dataset of the three candidates’ Twitter followers, we first analyzed the evolution in the composition of Sanders followers and then using neural network we studied whether there is a gender effect when Sanders supporters consider jumping ship for Trump.

Our study shows that between March and May an increasing proportion of Sanders followers are following Trump (but not Clinton). The proportion of Sanders followers who follow Clinton but not Trump has actually decreased. Equally important, our study suggests that the jumping ship behav-

⁸Washington Post, <https://www.washingtonpost.com/news/the-fix/wp/2016/05/24/how-likely-are-bernie-sanders-supporters-to-actually-vote-for-donald-trump-here-are-some-clues>.

ior will be affected by the gender effect and that men are more likely to switch to Trump than women.

6. ACKNOWLEDGMENT

We gratefully acknowledge support from the University and from our corporate sponsors Xerox and Yahoo.

7. REFERENCES

- [1] DOLAN, K. Is There a “Gender Affinity Effect” in American Politics? Information, Affect, and Candidate Sex in U.S. House Elections. *Political Research Quarterly* (2008).
- [2] FU, S., HE, H., AND HOU, Z.-G. Learning Race from Face: A Survey. In *Pattern Analysis and Machine Intelligence, IEEE Transactions on* (2014), vol. 36.
- [3] GINOSAR, S., RAKELLY, K., SACHS, S., YIN, B., AND EFROS, A. A. A Century of Portraits: A Visual Historical Record of American High School Yearbooks. In *ICCV 2015 Extreme Imaging Workshop Proceedings* (2015).
- [4] KING, D. C., AND MATLAND, R. E. Sex and the Grand Old Party: An Experimental Investigation of the Effect of Candidate Sex on Support for a Republican Candidate. *American Politics Research* (2003).
- [5] LEVI, G., AND HASSNER, T. Age and Gender Classification using Deep Convolutional Neural Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 34–42.
- [6] MACWILLIAMS, M. C. Forecasting Congressional Elections Using Facebook Data. *PS: Political Science & Politics* 48, 04 (October 2015).
- [7] O’CONNOR, B., BALASUBRAMANYAN, R., ROUTLEDGE, B. R., AND SMITH, N. A. From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media* (2010).
- [8] WANG, Y., FENG, Y., LI, Y., ZHANG, X., NIEMI, R., AND LUO, J. Pricing the Woman Card: Gender Politics between Hillary Clinton and Donald Trump. In *In arXiv preprint:1605.05401* (2016). In arXiv preprint:1605.05401 (2016).
- [9] WANG, Y., LI, Y., AND LUO, J. Deciphering the 2016 U.S. Presidential Campaign in the Twitter Sphere: A Comparison of the Trumpists and Clintonists. In *Tenth International AAAI Conference on Web and Social Media* (2016).
- [10] WANG, Y., LI, Y., YOU, Q., ZHANG, X., NIEMI, R., AND LUO, J. Voting with Feet: Who are Leaving Hillary Clinton and Donald Trump? In *arXiv preprint :1604.07103* (2016).
- [11] WANG, Y., LUO, J., NIEMI, R., AND LI, Y. To Follow or Not to Follow: Analyzing the Growth Patterns of the Trumpists on Twitter. In *Workshop Proceedings of the 10th International AAAI Conference on Web and Social Media* (2016).
- [12] WANG, Y., LUO, J., NIEMI, R., LI, Y., AND HU, T. Catching Fire via ‘Likes’: Inferring Topic Preferences of Trump Followers on Twitter. In *Tenth International AAAI Conference on Web and Social Media* (2016).