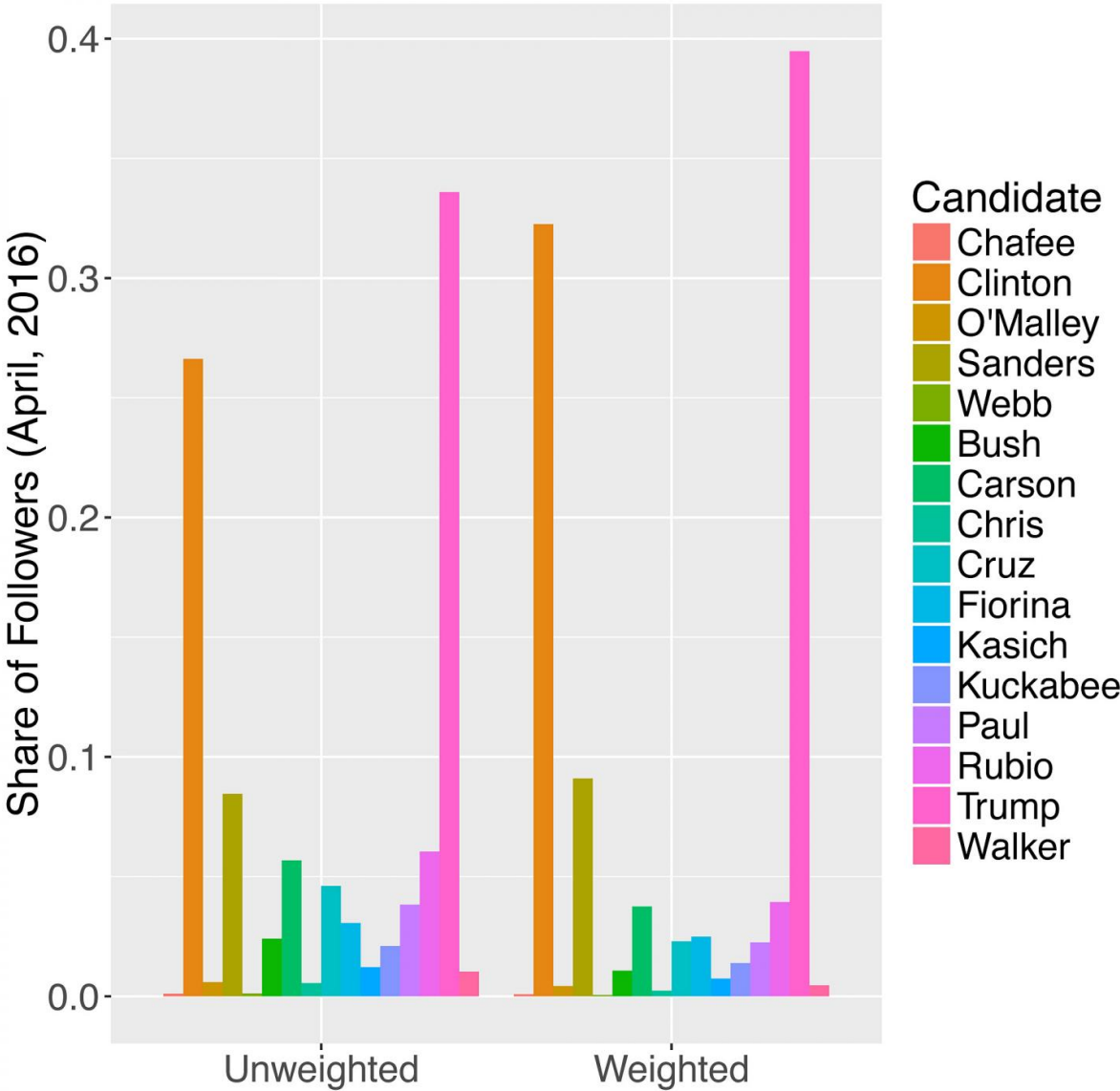


# Unlocking big data: Twitter researchers offer clues as to why Trump won

February 21 2017



The candidates' shares of total Twitter candidate followers in April 2016. The unweighted tallies simply count the number of followers. The weighted tallies take into account the fact that one individual can follow more than one candidate. As an example, an individual following two candidates has only a weight of 1/2, and an individual following three candidates has a weight of 1/3. By avoiding double counting, the weighted metric could better measure candidates' influence. Credit: Wang and Luo

Jiebo Luo and Yu Wang did not set out to predict who would win the 2016 U.S. presidential election. However, their exhaustive, 14-month study of each candidate's Twitter followers-enabled by machine learning and other data science tools-offers tantalizing clues as to why the race turned out the way it did.

"We wanted to understand how each of the candidate's campaigns evolved, and be able to explain why someone won or lost," says Luo, an associate professor of computer science.

Luo and Wang, a dual PhD candidate in political and computer science, summarized their findings in eight papers during the course of the campaign, including these observations:

— The more Donald Trump tweeted, the faster his following grew—even after he performed poorly in debates against other Republican candidates, and even after he sparked controversies, such as proposing a ban on Muslim immigration. (Read the paper at <https://arxiv.org/abs/1603.08174> )

— When Trump accused Hillary Clinton of playing the "woman card," women were more likely to follow Clinton and less likely to "un-follow" her during the week that followed. But it did not affect the gender composition of Trump followers. (Read the paper at

<https://arxiv.org/abs/1605.05401>)

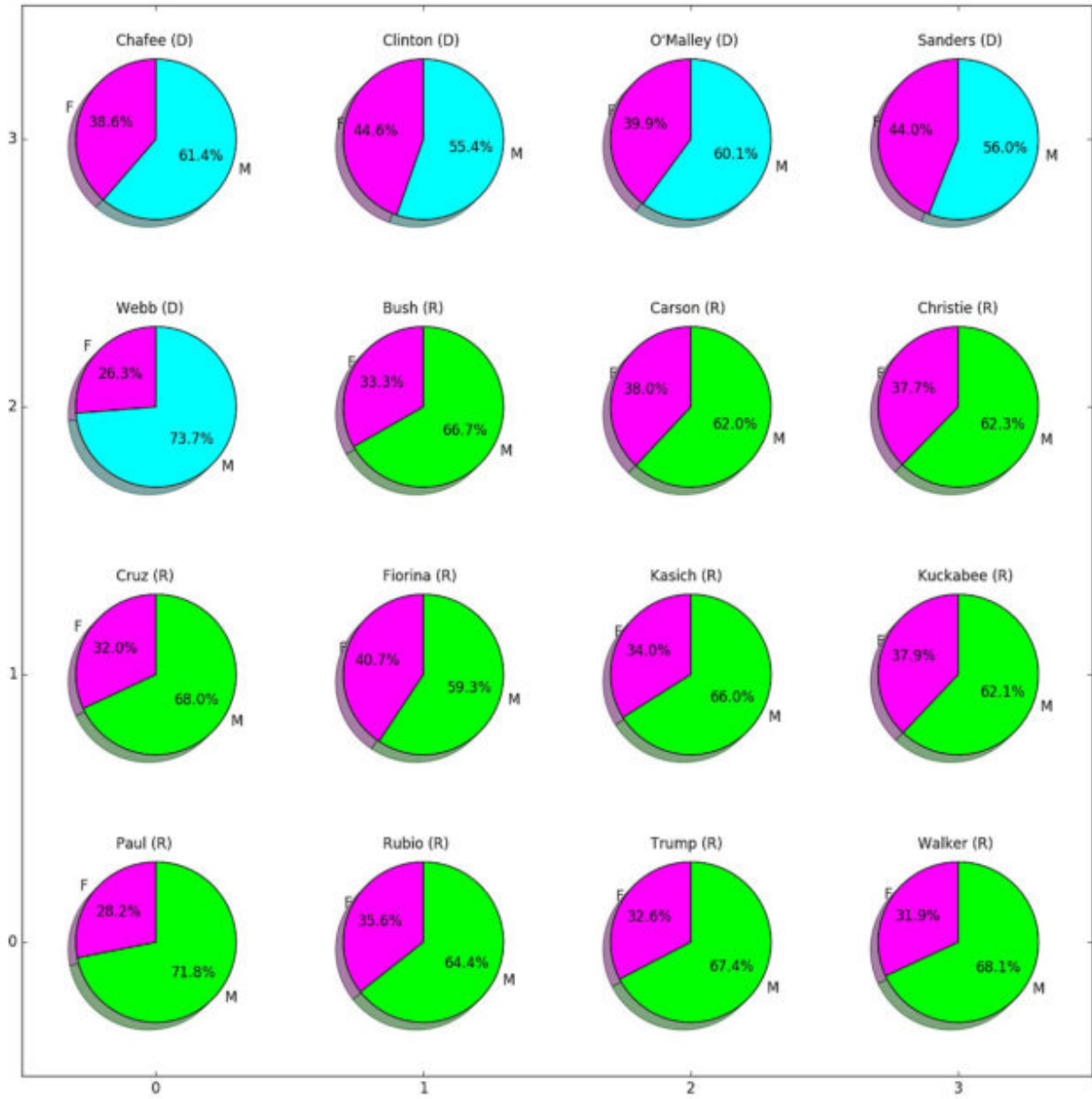
— Moreover, a "gender affinity effect" seen in other elections-women tending to vote for women-did not appear to be working for Clinton as the primaries drew to a close. The percentage of female Twitter followers in the Clinton camp was no larger than that in the Trump camp. Moreover, though "un-followers" were more likely to be female for both candidates, the phenomenon was "particularly pronounced" for Clinton. (Read the paper at <https://arxiv.org/abs/1604.07103> )

— At the same time, several polls, including ABC/Washington Post and CBS/New York Times, suggested that some Bernie Sanders supporters might "jump ship" from the Democratic column, and end up voting for Trump if Sanders dropped out. Luo and Wang found supporting evidence, reporting that the number of Bernie Sanders followers who were also following Trump was increasing-but the number also following Clinton was declining. The dual Sanders/Trump followers were also disproportionately (up to 64 percent) male. (Read the paper at <https://arxiv.org/abs/1605.09473> )

"In the end, even though we chose not to make any predictions, we were not surprised at all that Donald Trump won," says Luo.

## **Why Twitter?**

Barack Obama's use of social media in the 2008 presidential race helped establish Twitter and other social media platforms as powerful tools for candidates to quickly reach and receive feedback from large numbers of potential voters-and to attack their opponents.



Gender of candidate Twitter followers in April 2016, compiled by Wang and Luo. Credit: University of Rochester

Since then, there's been a burgeoning interest in scholarly research employing data science to analyze elections based on [social media](#) postings.

Twitter, in particular, is a rich source of data because the millions of tweets posted by its members each day are easily accessible using an application programming interface.

The key for Luo, Wang, and their colleagues was to collect as much of this data as possible, starting early in the campaign, and to then "mine" it in innovative ways.

"The very nature of this data is that it will disappear tomorrow, so we had to start capturing it from an early stage and design a research framework so we could continue to collect data all along," said Wang.

From September 2015 through October 2016, the team began accumulating a huge data set that included:

- The number of Twitter followers of each of the major candidates in the initially crowded field—updated every 10 minutes.
- 8 million tweets sampled from the followers of Clinton and Trump.
- 1 million images of the candidates' followers on Twitter.
- 5 million Twitter IDs that include all candidate followers in early April 2016.

Using advanced computer vision tools, the researchers trained an artificial neural network (what's called a convolutional neural network) to determine—with 90 percent accuracy or more—the age, gender, and race of the candidates' followers using their Twitter photos. This helped the researchers analyze the role of each of those factors in the campaign, as they tracked the changes in each candidate's followers before and after debates, for example, and how followers reacted to the candidates' own tweets.

Twitter mining has its limits compared to the responses gleaned from

traditional telephone polling. There's no opportunity to ask follow-up questions, for example, and tweets are difficult to place geographically, limiting their application for studying trends in swing states. (Even geotagged tweets may be sent while the sender is on vacation or attending a rally in another state.)

But Twitter mining also has its advantages-enabling researchers to quickly, continually, and inexpensively sample data on a scale that far surpasses the 1,000 or so responses that pollsters increasingly struggle to gather using traditional techniques. In one study, for example Luo and Wang were able to characterize 322,116 Trump or Clinton followers who subsequently became "un-[followers](#)."

"This is an approach that is broadly applicable," Luo says. "If you want to test public reaction to the next generation of iPhones, or to a new model of car, you can use the same approach to see what consumers like or don't like. It enables us to track millions of people and get reliable readings on their preferences."

Other Election 2016 papers by Luo, Wang, and their colleagues look at:

- [Gender Politics . . . A Computer Vision Approach](#)
- [Inferring Voter Preferences . . . Using Sparse Learning](#)
- [A Comparison of Trumpists and Clintonists](#)
- [Inferring Topic Preferences of Trump Followers](#)
- [Rumor Detection](#)
- [Election Bias: Comparing Polls and Twitter](#).

Provided by University of Rochester

Citation: Unlocking big data: Twitter researchers offer clues as to why Trump won (2017, February 21) retrieved 17 April 2024 from <https://techxplore.com/news/2017-02-big-twitter->

[clues-trump-won.html](https://clues-trump-won.html)

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.