

Finding Credibility Clues on Twitter

Millions of tweets analyzed to measure perceived trustworthiness

By Jason Maderer | January 26, 2017 • Atlanta, GA



By scanning 66 million tweets linked to nearly 1,400 real-world events, Georgia Institute of Technology researchers have built a language model that identifies words and phrases that lead to strong or weak perceived levels of credibility on Twitter. Their findings suggest that the words of millions of people on social media have considerable information about an event's credibility – even when an event is still ongoing.

“There have been many studies about social media credibility in recent years, but very little is known about what types of words or phrases create credibility perceptions during rapidly unfolding events,” said Tanushree Mitra, the Georgia Tech Ph.D. candidate who led the research.

The team looked at tweets surrounding events in 2014 and 2015, including the emergence of Ebola in West Africa, the Charlie Hebdo attack in Paris and the death of Eric Garner in New York City. They asked people to judge the posts on their credibility (from “certainly accurate” to “certainly inaccurate”). Then the team fed the words into a model that split them into 15 different linguistic categories. The classifications included positive and negative emotions, hedges and boosters, and anxiety.

The Georgia Tech computer then examined the words to judge if the tweets were credible or not. It matched the humans' opinions about 68 percent of the time. That's significantly higher than the random baseline of 25 percent.

“Tweets with booster words, such as ‘undeniable,’ and positive emotion terms, such as ‘eager’ and ‘terrific,’ were viewed as highly credible,” Mitra said. “Words indicating positive sentiment but mocking the impracticality of the



[Click image to enlarge](#)

Eric Gilbert and Tanushree Mitra

[MORE PHOTOS](#)

event, such as 'ha,' 'grins' or 'joking,' were seen as less credible. So were hedge words, including 'certain level' and 'suspects.'"

Higher numbers of retweets also correlated with lower credibility scores. Replies and retweets with longer message lengths were thought to be more credible.

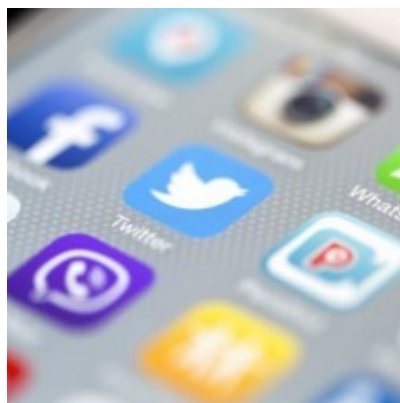
"It could be that longer message lengths provide more information or reasoning, so they're viewed as more trustworthy," she said. "On the other hand, a higher number of retweets, which was scored lower on credibility, might represent an attempt to elicit collective reasoning during times of crisis or uncertainty."

The system isn't deployable yet, but the Georgia Tech team says it could eventually become an app that displays the perceived trustworthiness of an event as it unfolds on social media.

"When combined with other signals, such as event topics or structural information, our linguistic result could be an important building block of an automated system," said Eric Gilbert, Mitra's advisor and an assistant professor in Georgia Tech's School of Interactive Computing. "Twitter is part of the problem with spreading untruthful news online. But it can also be part of the solution."

The paper, "A Parsimonious Language Model of Social Media Credibility Across Disparate Events," will be presented in February at the 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing in Portland, Oregon.

Additional Photos



Twitter and social media

Related Links

- Eric Gilbert's comp.social Lab

Contact Information

Jason Maderer
National Media Relations
maderer@gatech.edu
404-660-2926

Categories

Society and Culture