

The fake news detector

June 14 2018



There are many methods of determining if news is 'fake,' including analyzing the linguistic features of stories to extract key patterns. Credit: Patrick Dodson

A story on Reddit asks, "Did Palestinians Recognize Texas as Part of Mexico?" The origin of the story might be dubious, but it doesn't prevent the "fake news" story from accumulating 1.5 million likes across multiple platforms in just four days. The fake news dilemma dates back centuries, [according to Politico](#), but the advance of technology and the rise of social media, it's now at its zenith.

The problem of fake [news](#) fascinates Shivam Parikh, a doctoral student in UAlbany's College of Engineering and Applied Sciences. Parikh, working with Associate Professor of Computer Science Pradeep Atrey, recently presented on the subject at the IEEE 1st International

Conference on Multimedia Information Processing and Retrieval. Their paper, "Media-Rich Fake News Detection: A Survey," looks at the challenges associated with detecting fake news, existing detection approaches that are heavily based on text-based analysis, and popular fake news [data sets](#).

Fake news can be any content that is not truthful and is generated to convince its readers to believe in something that is not true, said Parikh, who works as a systems developer analyst for ITS at UAlbany.

The challenge for fake news detection comes with the democratization of news sources, and how easy modern technology makes sharing news articles in the age of social media.

Parikh and Atrey set out to address several critical pieces of the 'fake news' puzzle with their paper:

- The various platforms that can be used to disseminate content effectively and widely;
- The types of data news article may contain, and the impact of each type of data on readers;
- The different types of fake news categories;
- Existing fake news detection methods; and
- Current data sets that are available for fake news detection.

The researchers conclude by highlighting open research challenges in the area of fake news detection.

In 2017, two-thirds of U.S. adults get news from social media, a 5 percent jump over 2016, according to Reuters. Not surprisingly, this represents a blessing and a curse for the likes of Facebook and Twitter: the statistic represents the popularity of the platforms as well as their role as the primary sources for the spread of fake news.

But while the [social media](#) giants grapple with the misuse of their platforms, they are also confronted with the daunting nature of their task. Fake news can take on many forms, including photoshopped images, fake user-generated content or spoofed accounts, network-based content designed to appeal to a particular organization or group, as well as knowledge-based stories that contain a scientific or reasonable explanation of unresolved issues, often resulting in the spread of false information.

But while the task to detect fake news may sound daunting, there are several promising methods at researchers' disposal. Parikh and Atrey present a categorization of these approaches, their key characteristics and then analyze their respective advantages and limitations.

These methods include approaches that analyze linguistic features of stories to extract key patterns in fake news, or deception modeling, which is the process of clustering deceptive vs. truthful stories. Other approaches include developing predictive models that can assign positive or negative coefficients which can increase or decrease the probability of a story's truth; or content cue analysis, which is based on the ideology of what journalists like to write for users and what users like to read.

Fortunately, researchers have ample access to repositories of "fake news" articles in the form of publicly available data sets, such as BuzzFeedNews or LIAR. But while each of the data sets provides ample opportunity to study linguistic detection models, none possess a method for analyzing photos, for example.

"Visual presentation plays a huge role in people believing in fake news content. This calls for verification of not just language, but images, audio, embedded content, such as embedded video, tweet, Facebook post and hyperlinks," said Parikh.

Parikh also advocates for a detection method that can verify the source of the news story, and consider the trustworthiness or validity of the source once it's determined. An author credibility check can serve a similar function, where a system can be used to detect chains of [fake news](#) written by the same author or same group of authors.

Parikh knows he has only scratched the surface on the subject. Still, he is determined to explore the issue as he pursues his doctorate at UAlbany.

More information: Media-Rich Fake News Detection: A Survey.
www.albany.edu/~sp191221/publication/FakeNewsDetectionA_Survey.pdf

Provided by University at Albany

Citation: The fake news detector (2018, June 14) retrieved 10 April 2024 from
<https://phys.org/news/2018-06-fake-news-detector.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.
