Researchers identify seven types of fake news, aiding better detection
15 November 2019, by Matt Swayne

To help people spot fake news, or create technology that can automatically detect misleading content, scholars first need to know exactly what fake news is, according to a team of Penn State researchers. However, they add, that's not as simple as it sounds.

"There is a real crisis in our cultural understanding of the term 'fake news,'" said S. Shyam Sundar, James P. Jimirro Professor of Media Effects and co-director of the Media Effects Research Laboratory in the Donald P. Bellisario College of Communications.

In a study, researchers narrowed down myriad examples of fake news to seven basic categories, which include false news, polarized content, satire, misreporting, commentary, persuasive information and citizen journalism. The researchers also contrasted those types of content with real news and report their findings in the current issue of American Behavioral Scientist.

The researchers found that real news has message characteristics that differentiate it from the various categories of fake news, such as adherence to journalistic style. False news tends to be less grammatical and less factual, with greater reliance on emotionally charged claims, misleading headlines and so on. They also differ in the kinds of sources they use and how they use them.

In addition, the study noted differences in the structure of the site, such as the use of non-standard web addresses and personal e-mails in the "contact us" section. Furthermore, network differences can be used to help distinguish them, with fabricated news primarily circulated among social media accounts and seldom involving mainstream media outlets.

According to Maria Molina, a doctoral candidate in mass communications and lead author of the article, identifying the various message, source, structural and network features of different forms of online news is necessary to not just help people spot fake news, but to also help scientists who are using artificial intelligence—AI—to build systems that could one day automatically alert people to content that may be misinformation.

"In our own media environment we receive many different types of content, but not all of them are meant to inform. However, they all appear in the same format, so it is easy for people to confuse them with real news," said Molina. "And, in order to automatically detect fake news, we first need to understand exactly what fake news is and what the different layers are, so that we can classify one piece of content as fake compared to another piece of content."

The researchers used a research technique called a concept explication to undertake the study. The process requires researchers to conduct exhaustive searches of references to concepts, in this case, fake news, in scholarly and popular media. The researchers then examined how fake news is defined and how it is measured.
Online news content may also lack many of the structural cues once used by more traditional forms of media that helped people better differentiate between different forms of content. For example, commentary once appeared on the editorial section of a paper that signaled that the article was opinion. In addition, advertisements may have been set off in a box to separate it from news content, said Sundar, who is also an affiliate of Penn State's Institute for Computational and Data Sciences (ICDS), which provides Penn State faculty with supercomputing resources.

The researchers suggest that a better understanding of the various forms of fake and real news could lead to improved labeling of content, which could help restore some of that news segmentation. If content is properly labeled, online news consumers may have different reactions to different forms of news and information, according to Sundar.

"For example, if a piece of content is labeled as straight news, then it's a different story then if it's labeled commentary, or satire," he said. "So, we think it's very important to recognize the various elements of online news to be able to calibrate the expectations of readers and also of certain public figures who accuse the media of falsifying information."

Using computers to automatically detect fake news is difficult because these systems only see the content as either true or fake, said Dongwon Lee, the principal investigator of the project and an associate professor in the College of Information Sciences and Technology. Lee, who is also an affiliate of ICDS, said that is not always the case.

"As we encounter content in real life, the situation is much messier and murkier," said Lee. "For instance, despite containing some factually incorrect information, a satire article should not be blindly labeled as fake if the context is clear; yet, at the same time, if only some parts of the satire article are used, out of context, in social media, then it should be labeled as fake to curb its spread."

He added that this study's findings could be used to develop AI techniques that can identify multiple types of fake news, which will better reflect the real-world news environment.

"Our improved understanding in this article on the characteristics of seven sub-types on the spectrum of true-to-fake news will enable us to develop a new type of an auto-detection system capable of more fine-grained judgments," said Lee. "We are currently developing such a solution using the multinomial supervised learning technique in machine learning."


Provided by Pennsylvania State University