New AI tool tracks evolution of COVID-19 conspiracy theories on social media

April 19, 2021

LOS ALAMOS, N.M., April 19, 2021—A new machine-learning program accurately identifies COVID-19-related conspiracy theories on social media and models how they evolved over time—a tool that could someday help public health officials combat misinformation online.

"A lot of machine-learning studies related to misinformation on social media focus on identifying different kinds of conspiracy theories," said Courtney Shelley, a postdoctoral researcher in the Information Systems and Modeling Group at Los Alamos National Laboratory and co-author of the study that was published last week in the *Journal of Medical Internet Research.*

"Instead, we wanted to create a more cohesive understanding of how misinformation changes as it spreads. Because people tend to believe the first message they encounter, public health officials could someday monitor which conspiracy theories are gaining traction on social media and craft factual public information campaigns to preempt widespread acceptance of falsehoods."

The study, titled "Thought I'd Share First," used publicly available, anonymized Twitter data to characterize four COVID-19 conspiracy theory themes and provide context for each through the first five months of the pandemic.

The four themes the study examined were that 5G cell towers spread the virus; that the Bill and Melinda Gates Foundation engineered or has otherwise malicious intent related to COVID-19; that the virus was bioengineered or was developed in a laboratory; and that the COVID-19 vaccines, which were then all still in development, would be dangerous.

"We began with a dataset of approximately 1.8 million tweets that contained COVID-19 keywords or were from health-related Twitter accounts," said Dax Gerts, a computer scientist also in Los Alamos' Information Systems and Modeling Group and the study's co-author. "From this body of data, we identified subsets that matched the four conspiracy theories using pattern filtering, and hand labeled several hundred tweets in each conspiracy theory category to construct training sets."

Using the data collected for each of the four theories, the team built random forest machine-learning, or artificial intelligence (AI), models that categorized tweets as COVID-19 misinformation or not.

"This allowed us to observe the way individuals talk about these conspiracy theories on social media, and observe changes over time," said Gerts.

The study showed that misinformation tweets contain more negative sentiment when compared to factual tweets and that conspiracy theories evolve over time, incorporating details from unrelated conspiracy theories as well as real-world events.

For example, Bill Gates participated in a Reddit "Ask Me Anything" in March 2020, which highlighted Gates-funded research to develop injectable invisible ink that could be used to record vaccinations. Immediately after, there was an increase in the prominence of words associated with vaccine-averse conspiracy theories suggesting the COVID-19 vaccine would secretly microchip individuals for population control.

Furthermore, the study found that a supervised learning technique could be used to automatically identify conspiracy theories, and that an unsupervised learning approach (dynamic topic modeling) could be used to explore changes in word importance among topics within each theory.

"It's important for public health officials to know how conspiracy theories are evolving and gaining traction over time," said Shelley. "If not, they run the risk of inadvertently publicizing conspiracy theories that might otherwise 'die on the vine.' So, knowing how conspiracy theories are changing and perhaps incorporating other theories or real-world events is important when strategizing how to counter them with factual public information campaigns."

Unable to load contents of IFRAME at this location in the original document. See original HTML document and notify an administrator.

Interview with Ashlynn Daughton, an information scientist in the Information Systems and Modeling Group at Los Alamos National Laboratory, and co-author of the study.

LA-UR-21-23512

Los Alamos National Laboratory   www.lanl.gov   (505) 667-7000   Los Alamos, NM

Managed by Triad National Security, LLC for the U.S Department of Energy's NNSA   **NNSA** National Nuclear Security Administration