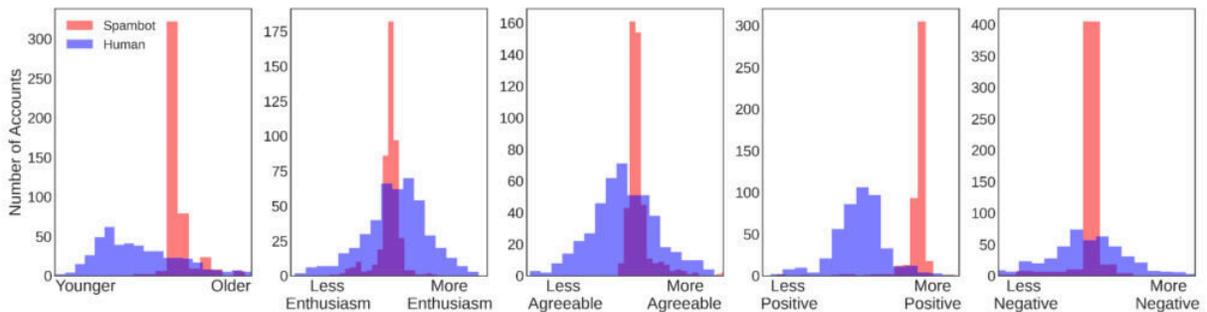


Social media bots may appear human, but their similar personalities give them away

November 24 2021, by Evan Lerner



Distribution of human (blue) and bot (red) accounts across age, personality, and sentiment. Across each trait, the human accounts have a large spread of values, whereas the bot accounts are all clustered within a small range. For enthusiasm, agreeableness, and negativity the bots are clustered around the center of the human distribution, showing that these accounts exhibit very average traits. Credit: University of Pennsylvania

Social bots, or automated social media accounts that pose as genuine people, have infiltrated all manner of discussions, including conversations about consequential topics, such as the COVID-19 pandemic. These bots are not like robocalls or spam emails; recent studies have shown that social media users find them mostly indistinguishable from real humans.

Now, a new study by University of Pennsylvania and Stony Brook

University researchers, published in Findings of the Association for Computational Linguistics, gives a closer look at how these bots disguise themselves. Through state-of-the-art machine learning and [natural language](#) processing techniques, the researchers estimated how well bots mimic 17 human attributes, including age, gender, and a range of emotions.

The study sheds light on how bots behave on [social media platforms](#) and interact with genuine accounts, as well as the current capabilities of bot-generation technologies.

It also suggests a new strategy for detecting bots: While the language used by any one bot reflected convincingly human [personality traits](#), their similarity to one another betrayed their artificial nature.

"This research gives us insight into how bots are able to engage with these platforms undetected," said lead author Salvatore Giorgi, a graduate student in the Department of Computer and Information Science (CIS) in Penn's School of Engineering and Applied Science. "If a Twitter user thinks an account is human, then they may be more likely to engage with that account. Depending on the bot's intent, the end result of this interaction could be innocuous, but it could also lead to engaging with potentially dangerous misinformation."

Along with Giorgi, the research was conducted by Lyle Ungar, Professor in CIS, and senior author H. Andrew Schwartz, Associate Professor in the Department of Computer Science at Stony Brook University.

Some of the researchers' previous work showed how the language of social media posts can be used to accurately predict a number of attributes of the author, including their age, gender and how they would score on a test of the "Big Five" personality traits: openness to experience, conscientiousness, extraversion, agreeableness and

neuroticism.

The new study looked at more than 3 million tweets authored by three thousand bot accounts and an equal number of genuine accounts. Based only on the language from these tweets, the researchers estimated 17 features for each [account](#): age, gender, the Big Five personality traits, eight emotions (such as joy, anger and fear), and positive/negative sentiment.

Their results showed that, individually, the bots look human, having reasonable values for their estimated demographics, emotions and personality traits. However, as a whole, the social bots look like clones of one another, in terms of their estimated values across all 17 attributes.

Overwhelmingly, the language bots used appeared to be characteristic of a person in their late 20s and overwhelmingly positive.

The researchers' analysis revealed that the uniformity of social bots' scores on the 17 human traits was so strong that they decided to test how well these traits would work as the only inputs to a bot detector.

"Imagine you're trying to find spies in a crowd, all with very good but also very similar disguises," says Schwartz. "Looking at each one individually, they look authentic and blend in extremely well. However, when you zoom out and look at the entire crowd, they are obvious because the disguise is just so common. The way we interact with social media, we are not zoomed out, we just see a few messages at once. This approach gives researchers and security analysts a big picture view to better see the common disguise of the social bots."

Typically, bot detectors rely on more features or a complex combination of information from the bot's social network and images they post. Schwartz and Giorgi found that by automatically clustering the accounts

into two groups based only on these 17 traits and no bot labels, one of the two groups ended up being almost entirely bots. In fact, they were able to use this technique to make an unsupervised bot detector, and it approximately matched state-of-art-accuracy (true-positive rate: 0.99, sensitivity/recall: 0.95).

"The results were not at all what we expected," Giorgi said. "The initial hypothesis was that the social bot accounts would clearly look inhuman. For example, we thought our classifier might estimate a bot's age to be 130 or negative 50, meaning that a real user would be able to tell that something was off. But across all 17 traits we mostly found bots fell within a 'human' range, even though there was extremely limited variation across the bot population."

By contrast, when looking at the human trait estimates of non-social bots, automated accounts that make no attempt to hide their artificial nature, the trait distributions matched those of the original hypothesis: estimated values that fell outside of normal human ranges and seemingly random distributions at the population level.

"There is a lot of variation in the type of accounts one can encounter on Twitter, with an almost science fiction-like landscape: humans, human-like clones pretending to be humans, and robots," says Giorgi.

More information: Salvatore Giorgi et al, Characterizing Social Spambots by their Human Traits, *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021* (2021). [DOI: 10.18653/v1/2021.findings-acl.457](https://doi.org/10.18653/v1/2021.findings-acl.457)

Provided by University of Pennsylvania

Citation: Social media bots may appear human, but their similar personalities give them away (2021, November 24) retrieved 24 April 2024 from <https://techxplore.com/news/2021-11-social-media-bots-human-similar.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.