Friends of friends reveal our hidden online traits

April 6 2018, by Tom Abate

Credit: CC0 Public Domain

At a time when social network privacy – or the lack of it – is headline news, two Stanford researchers have some sobering findings about how personal data is becoming increasingly difficult to hide if we have any

public presence online.

In a paper published this month in *Nature Human Behavior*, Johan Ugander, assistant professor of management science and engineering, and Kristen Altenburger, a Ph.D. student in his lab, have shown that there are more ways than previously realized to reveal demographic traits that people might be trying to conceal. This work builds on one of the main threads in privacy research, which is to understand how different traits are correlated.

The Stanford paper is based on databases made available specifically for research. These reflect the kinds of information that websites make available to advertisers or reveal to outside groups when people allow third parties to access their social profiles. Given the prevalence of such data, the researchers sought to better understand what sorts of statistical inferences might end up revealing traits people have sought to conceal.

"In social data, some things are more predictable than others," Ugander said. "We set out to study the relationship between friend networks and predictability, and ended up uncovering an inference mechanism that hadn't been noticed before."

Levels of inference

At the simplest level people reveal information about themselves based on how they behave online. If a person buys diapers online, for example, they probably have a baby. That is a direct inference.

A second form of inference is based on looking at our friends, or indirect inference. Researchers who have studied social media relationships have found that we tend to friend people of roughly our same age, race and political belief. So even if a person does not reveal their age, race or political views, these traits are easily and accurately

inferred from friendship studies. Researchers call this tendency homophily, which stems from the Greek words for love of sameness.

But not all unknown traits are easy to predict using friend studies. Gender, for instance, exhibits what researchers call weak homophily in online contexts.

"If an unknown person in a social network has mostly male friends there's an almost equally good chance they could be female, or vice versa," Altenburger said.

Friends of friends

The group's new research shows that it's possible to infer certain concealed traits – gender being the first – by studying the friends of our friends.

This technique works because Ugander and Altenburger have described a new social structure they call monophily, Greek for "love of one," where people have extreme preferences for traits but not necessarily their own trait. "For example," Ugander said, "on average it might be the case that men don't have a clear preference for male or female friends, but that average may be obscuring the fact that some men have strong preferences for male friends while others have strong preferences for female friends."

They observe that when there's monophily in a network, it becomes possible to predict traits of individuals based on friends of friends, even in situations where there's no homophily.

The Stanford team relied on standard network datasets widely studied by academics. These datasets map friendship networks and contain complete information about all of the traits of all of the individual traits,

including gender. The researchers then erased the gender data for certain individuals, creating artificial unknowns, and then used their "friends of friends" analysis to see if it could make a prediction.

"It's a fill-in-the-blanks problem," said Ugander. "And while we find that your friends don't tend to predict your gender, the people those friends choose to associate with, your friends of friends, tend to be more similar to you than even your friends are."

The researchers said that the power of their new perspective, of looking at the friends of our friends, highlights the importance of protecting network data from prying hands. Any policy solution to preserve network privacy will need to consider the information contained among one's friends of friends. They are now reapplying their technique to other unknowns to see what else may be disclosed by friends of friends.

"We're not sure what else might be revealed in this way," Ugander said, adding: "Unfortunately, it looks like the realm of [network] privacy is even smaller than we previously thought."

More information: Kristen M. Altenburger et al. Monophily in social networks introduces similarity among friends-of-friends, *Nature Human Behaviour* (2018). [DOI: 10.1038/s41562-018-0321-8](https://doi.org/10.1038/s41562-018-0321-8)

Provided by Stanford University

provided for information purposes only.