You are what you click

September 29 2015, by Sonia Fernandez

Credit: MATT PERKO

It's no secret that the things we click on, scroll across, swipe, tap or drag when we're browsing online or using a smartphone application can yield valuable information about us. Such data is a veritable goldmine to web browsers and online retailers who use it to assess our preferences and target advertising to our tastes. But, researchers at UC Santa Barbara

suggest that studying users' online or smartphone actions could yield far more information about us than simply shopping habits.

With $499,949 in funding from the National Science Foundation, computer science professors Ben Zhao and Heather Zheng will investigate how these detailed, time-stamped logs of online and smartphone user actions, dubbed "clickstreams," can reveal insights about users.

"We're trying to build natural, data-driven models of user behavior, without external labels or preconceived categories, but rather from natural patterns in user behavior," said Zhao. To do this, the researchers will analyze anonymized user action logs from Los Angeles-based smartphone app Whisper and Chinese social media site Renren. "By avoiding preset categories, we can identify real natural patterns that may be difficult to predict, while allowing for real unexpected surprises in user behavior," he explained.

Using similarity graphs to compare the data from the multitudes of clickstreams, the scientists hope to identify general behavior patterns over time. This information can in turn be used to identify abnormal or unusual behaviors that could, for instance, indicate security risks.

"We've already done this successfully on LinkedIn and Renren to identify users using fake accounts to attack those systems," said Zhao. "The resulting accuracy has been exceptional, and has detected attacks that we could not have predicted in advance, simply by noting how large groups of users act similarly to each other but unlike most normal users in the system."

In the big picture, the researchers will also be investigating methods of making these clickstream graphs scalable to large user populations—in the hundreds of millions of users—as well as ways to tune the modeling

system to different dimensions of user behavior or to specific applications.

Provided by University of California - Santa Barbara