Senator Marsha Blackburn Questions
Senate Judiciary Subcommittee on Privacy, Technology, and the Law Hearing
"Protecting Americans' Private Data from Hostile Foreign Powers"
Wednesday, September 14, 2022
4:00 PM

Question for Professor Landau and Mr. Pottinger: I understand that foreign countries and companies under state ownership or control can use click through ads to gain access to platform user data. For example, in places like China, where Twitter is blocked, the CCP could use ads on the platform to collect data on users evading the block through virtual private networks.

A. *In your experience, is this a typical practice that happens at global tech platforms?*

B. *What types of data can be collected through such a process and what kinds of harms could it cause?*

A. Yes, this is a typical practice. The business model for many platforms is based on the harvesting and sale of user information and metadata, typically through the resale of information via data brokers. These business practices are fundamental to the revenue model of the platforms. Although the typical design of these platforms is based on aggregation of user data to form behavioral profiles based on marketing objectives (selling goods for profit) it is entirely possible to single out specific user data and information using metadata such as their location, prior behavioral activities and even their screen configuration or apps loaded to their machine. The success of these data collection methods varies depending on the privacy profile of the user, the technologies employed by the platform, and the technical means employed to facilitate data collection.

B. The data that can be collected through these means, depending on the technologies employed and the privacy settings of the targeted individual, can be extensive. Generally, the information collected is used to group people by their location and behavioral profiles (usually mapped to a profiling system) with the intention of predicting and thereby influencing behavior. However, if a specific individual is selected for detailed analysis, specific location and pattern of life data can be interrogated, browsing history can be reviewed, and networked graphs can be used to consider the connections between people, their interests, and pattern of life.

   Although browsers and apps are the main way of collecting information, the broader collection methods available all contribute to a larger picture that supports identity attribution. Machines can be fingerprinted based on their screen resolution, operating system, and other attributable information. Clipboard information can be accessed by some apps and programs. Website requests can be eavesdropped. Although some modern phones offer a degree of protection through virtual sandboxes, these do not offer protections against much of the data collection undertaken.

   The proliferation of low cost or free VPNs which are owned by private firms who are also on-selling data means that only savvy users are likely to reduce their overall fingerprint, and even then in most cases they can be identified and tracked online through targeted means. The biggest data collection opportunities come through apps and advertisements on browsers when users interact with them, but there are a variety of other techniques that can be employed. Combined with breached information available on the darknet, extensive datasets can be collated to link virtual with physical identities.

   In terms of harms caused, it is entirely possible to locate, surveil and target dissidents, protestors, journalists and even politicians with information which could be used to coerce or intimidate. It is also possible that potential target audiences for future influence activities could be identified to undermine the fabric of democratic processes and create dissent within the community, based on online behavioral profiles.