MIT
Technology
Review

Mobile Call Logs Can Reveal a Lot to the NSA

By Jessica Leber – June 18th, 2013

Of all the recent revelations about the National Security Agency's sweeping surveillance activities, the collection of metadata from Verizon's U.S. call records may be the most concerning (see "NSA Surveillance Reflects a Broader Interpretation of the Patriot Act"). Despite reassurances that the information collected is limited in its scope, academics who study such data say it could still reveal a great deal about the people being monitored.

To defend the program, members of Congress have been instructed to refer to these logs as mere "bare-bones records," according to a set of leaked talking points. But in reality, the metadata subject to the court order obtained by the NSA—including phone numbers, call time and duration, and information about device interactions with cellular towers—gives intelligence analysts a clear window into sensitive interactions and movements of the U.S. population.

The term "metadata" simply refers to information that is used to track or describe another piece of data, whether that is a cell-phone conversation or a money transfer. One study published this March, using records provided by a European wireless carrier, shows the surveillance power of telecommunications metadata. Vincent Blondel, an applied mathematician at MIT and the Université Catholique de Louvain in Belgium, and collaborators analyzed 15 months of anonymous call records from 1.5 million people. His team was able to uniquely pinpoint the movements of 95 percent of people from only four records, using only the location of a nearby cellular station and the time each call was made.

"You can infer a lot, such as where people work and where people live," says Blondel. "You don't need information about the content [of the call]."

From there, connecting such movements with a person's real identity would be a relatively simple matter of cross-referencing the records with other sources of data. The NSA may be able to do this using credit card transactions or e-mail communications, Blondel says, or simply by knowing who has the phone number.

These techniques could reveal sensitive activity such as attendance at a particular church or a visit to an abortion clinic. Analysts could even surmise where and when two individuals are meeting face to face, Blondel says, or construct a diagram of complex relationships among far-flung communities (see "Has Big Data Made Anonymity Impossible?").

It's not known exactly how and to what extent the NSA is mining Verizon's data, other than that it is collecting the information every day. And the practice may well assist the NSA in locating terrorism suspects or networks that are actively seeking to evade detection. "These individuals are very aware that there is a high probability their communications are being monitored," says Drew Conway, a scientist in residence at IA Ventures, who has studied data mining techniques used in counterterrorism efforts. But a more computationally intensive "big data" approach, without specific targets or questions in mind, could also be used to flag unusual communication patterns in hopes of predicting suspicious activity, he says.

Similar tactics have been used successfully against the U.S. In 2011, the Lebanese militant party Hezbollah was reportedly able to pinpoint a CIA network by mining cell-phone data for anomalies, like phones that were used only rarely and in specific locations.

In recent years, some mobile carriers have made various kinds of anonymous metadata available for uses ranging from marketing (see "How Wireless Carriers Are Monetizing Your Movements") to development studies and transportation planning (see "Glimpses of a World Revealed by Cell-Phone Data"). Unlike what the NSA is collecting, these data sets don't include actual phone numbers or other unique personal identifiers. However, the methods of mining and gleaning information are similar, and so are its potential uses.

In the aftermath of this month's revelations, privacy advocates are worried about what other kinds of metadata the NSA may collect now or in the future. The *Wall Street Journal* has reported, for example, that credit card purchases and Internet service provider (ISP) metadata may be part of a similar program, along with call data from AT&T and Sprint. Technology companies like Google and Facebook have denied that they hand over metadata on a comprehensive basis, but they have also said they fear an order to do so. What, exactly, metadata means is fuzzier online than in the context of phone calls. If this type of data were gathered, it would deeply extend the reach of the government's surveillance.

"One of the problems here is that metadata is kind of a relative term," says Julian Sanchez, a research fellow at the Cato Institute in Washington, D.C. "There's information that is metadata to Facebook and Google that might be data to the ISP."

In the context of e-mail, for example, metadata might mean the IP addresses of the sender and recipient and perhaps a time stamp or, arguably, the subject line. For Facebook, perhaps it's information about when a friend request is being made.

More broadly, whether the security benefits of this sweeping surveillance scheme outweigh the privacy costs won't be clear unless the Obama administration declassifies detailed information about instances in which such data proved crucial to combating terrorism but would have been difficult to obtain with a warrant. And Sanchez believes it is unlikely the government will prove its case.

"It may well be the case that this kind of thing is of some utility in some situations," he says, but only "in the same way a general warrant to search any house you please might be useful in preventing crime."