Technology Review
PUBLISHED BY MIT

Wednesday, July 28, 2010

The Hunt for the Wikileaks Whistle-blower

Digital encoding could catch future informants.

By David Talbot

Attorney General Eric Holder's new probe into Wikileaks's (http://www.wikileaks.org/) posting of 91,000 war documents will likely find that tracing the path of the documents back through the Internet is next to impossible. But watermarks--if they were embedded in the files--could reveal the whistle-blower.

Wikileaks relies on a networking technology called Tor (http://www.torproject.org/) , which obscures the source of uploaded data (http://www.technologyreview.com/computing/22427/?a=f) . While Tor doesn't encrypt the underlying data--that's up to the user--it does bounce the data through multiple nodes. At each step, it encrypts the network address. The source of data can be traced to the last node (the so-called "exit node"), but that node won't bear any relationship to the original sender.

Ethan Zuckerman, cofounder of the blogging advocacy organization Global Voices (http://globalvoicesonline.org/) , says he doubts investigators can crack Tor to find the computer from which the documents were originally sent. "There's been an enormous amount of research done on the security of the Tor network and on the basic security of encryption protocols," he says. "There are theoretical attacks on Tor that have been demonstrated to work in the lab, but no credible field reports of Tor being broken."

And while Tor's profile has been raised by its association with Wikileaks, Andrew Lewman, Tor's executive director, says he has no insights into the source of the purloined documents. "I don't know how Wikileaks got any of the information," he says. While Wikileaks gets technical help from Tor staffers, "they don't tell us anything, other than 'Did we set up the hidden service correctly?' which we'd answer for anyone," Lewman adds.

"People assume that Wikileaks is a Tor project, but I can tell you definitely there is no official relationship."

Lewman points out that many law-enforcement agencies, such as the U.S. Drug Enforcement Agency, also use Tor to protect their operations.

One way the government could finger a leaker is through digital watermarking of the documents themselves. James Goldman (http://www2.tech.purdue.edu/cpt/SelfStudy/CPTFacultyVitas/FacultyStaff/DisplayStaffMember.asp?member=jegoldman) , a cyber forensics expert at Purdue University, says it's not clear whether the government uses digital watermarking, "but it's certainly possible."

Such watermarks would consist of hidden digital data--or even slight alterations in the pattern of words--added to documents in ways that are hard to detect, but are readily decodable with the right software.

"If I'm in the government and charged with plugging a hole or catching a leak in the next 10 minutes, my attention is going to turn to watermarking," says Jonathan Zittrain (http://cyber.law.harvard.edu/people/jzittrain) , founder of the Berkman Center for Internet and Society (http://cyber.law.harvard.edu/) at Harvard University, and an Internet law professor there. "It wouldn't take much effort within the government to personalize a document to identify its recipient," so that this person could be identified if they later leaked that document.

Zuckerman adds that it's also probably safe to say that the basic cryptography that's widely used on the Internet--automatically deployed on banking websites and others via Web addresses that start with "https"--is also fairly secure. "It's impossible to say whether [the National Security Agency] has broken them, but most people who aren't unhealthily paranoid tend to believe that if [encryption] were badly broken ... we'd see theft of credit-card information on a massive scale."

While the outcome of Holder's investigation is hard to predict, it's a safe bet that the saga will result in an overhaul of how the government protects information. In addition to using watermarking, government agencies could adapt existing digital-rights-management technologies.

Such technologies can perform various tasks that might be relevant: generally, they can identify when the same computer is downloading voluminous amounts of material, restrict downloading to authorized users, and stop users from copying or passing restricted files to other computers. For example, a song purchased and downloaded onto one iPod in a protected format cannot easily and legally be transferred to other iPods.

"If you think about the technology of digital-rights management: How is it that the recording industry is able to hang on to the stuff in a way that the military can't?" says John Pike, director of Global Security.org (http://www.globalsecurity.org/) , the national security think tank. "It's hard to understand."

Upcoming Events

2010 IEEE Conference on Innovative Technologies for an Efficient and Reliable Electricity Supply (http://www.ieee-energy.org/)
Waltham, Massachusetts
Sunday, September 27, 2009 - Tuesday, September 28, 2010
http://www.ieee-energy.org/ (http://www.ieee-energy.org/)

FutureM (http://www.futurem.org/Default.aspx)
Boston, MA
Monday, October 04, 2010 - Friday, October 08, 2010
http://www.futurem.org/Default.aspx (http://www.futurem.org/Default.aspx)

USA Science & Engineering Festival Expo (http://www.usasciencefestival.org)
Washington, D.C.
Saturday, October 23, 2010 - Friday, October 29, 2010
http://www.usasciencefestival.org (http://www.usasciencefestival.org)

15th Annual MITX Interactive Awards (http://www.mitxawards.org/interactive/default.aspx)
Boston, Massachusetts
Thursday, November 18, 2010
http://www.mitxawards.org/interactive/default.aspx (http://www.mitxawards.org/interactive/default.aspx)