washingtonpost.com
Hello wehate.mandatory
Change Preferences | Sign Out
The Washington Post
Print Edition | Subscribe

NEWS | OPINIONS | SPORTS | ARTS & LIVING | DISCUSSIONS | PHOTOS & VIDEO | CITY GUIDE | CLASSIFIEDS | JOBS | CARS | REAL ESTATE | Shopping Deals »

SEARCH: go ● News ○ Web results by Google™ | Search Archives

Real Estate: Buy | Sell | Recent Sales | Directory
Rentals: Find a Rental | Rent Your Place | Directory

washingtonpost.com > Technology > Special Reports > Privacy

TechNews.com
Print This Article
E-Mail This Article

QUICK QUOTES
Enter Symbol go
Tables | Portfolio | Index

MOST VIEWED ARTICLES
Technology On the Site
Updated 8:16 a.m. ET
- Thieves Track Down GPS Units
- Ringing Up Big Music Sales
- NSA Program Further Blurs Line on Privacy
- Microsoft sides with Nintendo in fight vs Sony
- Nintendo, Activision seen benefiting from new consoles

RSS NEWS FEEDS
Top News
Technology
What is RSS? | All RSS Feeds

E-MAIL NEWSLETTERS
View a Sample and Sign Up
TechNews Daily Report
Personal Finance
Personal Tech

Manage Your Newsletters

big brother alert

Phone Calls Are Just the Start

By Simson L. Garfinkel
Sunday, May 14, 2006; Page B02

On Thursday, news accounts revealed that AT&T, Verizon and BellSouth have been sharing the telephone records of millions of Americans with the National Security Agency. According to USA Today, the NSA is using this information to create a database of every phone call being made within the nation's borders. The spy agency would then mine this database to uncover hidden terrorist networks. "It's the largest database ever assembled," a source told the newspaper.

The news spurred outrage on Capitol Hill. However, telephone records are just a sliver of the data on individuals that the government could assemble. Through our movements, transactions and activities, residents of industrialized societies throw off megabytes of data each day. Gathering this data is technically straightforward, and the potential for authorities to build much larger databases -- relying on sources we may not have contemplated before -- is quite real. Such databases would require extensive protections to prevent abuse from low-level insiders and senior government officials.

Building these databases can be incredibly time-consuming and expensive, and doing something useful with that data -- turning countless facts into actionable information -- is even more complicated. Data mining on a broad scale may indeed help the government in its professed goal of fighting terrorism. But are the costs involved and the risks to privacy from data-collection "mission creep" worth the effort?

⊞ Enlarge This Photo Buy This Photo

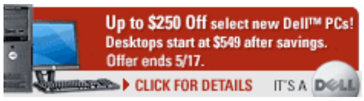
Much more to mine: Airline records, ATM transactions and dining reservations could be next on the government's surveillance list. (By Rebecca Mcalpin -- Bloomberg)

Government authorities can turn to many sources to feed their data-mining systems. Wireless telephone providers such as Verizon, for example, know a lot more about their customers than what numbers

PRIVACY

An entire industry has mushroomed during the past decade because of the ability of companies to gather and make sense of public records, criminal histories and other electronic details. What are they doing with it?

Lawyer: Ex-Qwest Exec Ignored NSA Request
GOP Duo Back Hayden for CIA
Verizon: No unfettered access for govt
Poll: Most Americans Support NSA's Efforts
White House Stands by Hayden Nomination



they dial. Cellular phones aren't just communication devices -- they're tracking devices, too. To route telephone calls, your cellphone provider must know your location whenever your phone is on. (According to a BBC report, geo-location data were used last July to track and stop a would-be suicide bomber in London.) Instead of having a watch list for individuals, the Department of Homeland Security could create an electronic watch for specific locations. A cellphone near the perimeter of a nuclear power plant might generate a database record; three such records in one year might trigger an investigation.

More News

SAVE & SHARE

Tag This Article

▸ Saving options

Powered by Del.icio.us

Such geo-location information could also be stored and used retrospectively for any number of purposes. If intelligence officers discover that terrorists held a planning meeting in some remote corner of a park, they could consult the database to see which phones were there. Authorities also could compile an electronic list of known associates by looking for cellphones that tend to travel in the same car or bus as a suspect.

Of course, data surveillance doesn't depend on cellphones -- it depends on data. If Alice is on a terrorist watch list, the government may decide it has an interest in everybody who calls Alice regularly. If Bob is one of those frequent callers, the government may decide that he should be under surveillance, too.

And other kinds of everyday information could help locate even tech-savvy terrorists who try hard to avoid leaving an electronic footprint. For example, a database of airline reservations could reveal that Alice and Bob frequently travel on the same flight, or that they travel to the same cities on the same days from different locations. Credit card charges or ATM withdrawals or even records from prepaid calling cards can provide similar information. Such data has already proved useful: The FBI used records from prepaid calling cards to find a close link between Oklahoma City bombers Timothy McVeigh and Terry Nichols. This kind of social-network analysis requires a lot of data and very fast computers -- and neither of those is in short supply.

But although it's straightforward to identify social networks, it's much harder to find social networks that are meaningful. Alice and Bob might eat at the same restaurants on the same nights because they are having covert meetings -- or because they are both fans of a local rock band. They might even follow the band to different cities, potentially tripping another alert in some government computer model.

Indeed, the real danger of this kind of unrestricted data mining isn't "false positives" -- that is, associations that don't really exist -- but meaningless positives. Investigating all of these positives takes time and money. And if the investigations are not done correctly, innocent lives can be ruined in the process. A principle of American jurisprudence is that it is better for a guilty person to go free than for an innocent person to be imprisoned. How will our society react to a system that requires many people to be investigated because only one of them might be a terrorist?

Of course, critics should recognize that large-scale data surveillance can do much more than simply find new terrorism suspects; it could detect the outbreak of a disease days before medical professionals become aware of it. The Realtime Outbreak and Disease Surveillance project at the University of Pittsburgh monitors the sale of over-the-counter cold remedies and other health-care products at more than 20,000 stores throughout the United States. Its purpose: Identify disease outbreaks early, while there is still a chance to contain them. People who get sick will try self-medicating before seeking professional help. A compelling graphic on the RODS Web site shows that over-the-counter sales of cold medications can peak two weeks before hospital admissions.

The RODS researchers stress that their bio-surveillance does not violate privacy because no personally identifiable information is ever assembled or reported. The system just collects aggregate sales from a sampling of the nation's largest drugstores and mass-merchandise chains. In the case of an actual bioterrorism attack, it would be helpful to know the names of the people who were infected. But collecting that information isn't necessary to achieve the project's primary goals, and it would create unacceptable risks to those involved because the data would be so easily subject to abuse.

Ultimately, that may be the greatest dilemma for those involved in collecting and mining data: What information does one *not* need to collect, and when is it safe to throw away a piece of data? It's human nature to hold on to information as long as possible -- once you eliminate it, you can't always get it back. And even if you could keep everything forever, would you want to? The cost of storing and protecting data is high. The more information you have, the more difficult it is to search and cross-reference it, which means you must spend more on computer systems. And perhaps the most insidious side effect of all: After spending all the money and effort to collect, keep and protect data, it is hard not to develop that nagging feeling that you really should be putting it to use.

Simson L. Garfinkel is a postdoctoral fellow at Harvard University's Center for Research on Computation and Society and the author of "Database Nation: The Death of Privacy in the 21st Century" (O'Reilly).

MORE TECHNOLOGY ARTICLES

Most Viewed Technology Articles

• Thieves Track Down GPS Units
• Ringing Up Big Music Sales
• NSA Program Further Blurs Line on Privacy
• Microsoft sides with Nintendo in fight vs Sony
• Nintendo, Activision seen benefiting from new consoles

» **Top 35 Most Viewed**

Editors' Picks: Technology

• Ringing Up Big Music Sales
• NSA Program Further Blurs Line on Privacy
• India's Ragtag Band of Maoists Takes Root Among Rural Poor
• Thieves Track Down GPS Units
• Photo Blogging

» **More in the Technology Section**

Print This Article E-Mail This Article

SEARCH: go ● News ○ Web results by Google™

NEWS | OPINIONS | SPORTS | ARTS & LIVING Discussions | Photos & Video | City Guide CLASSIFIEDS | JOBS | CARS | REAL ESTATE

washingtonpost.com: Help | Contact Us | About Us | Advertise With Us | Site Index | Site Map | Make Us Your Homepage | mywashingtonpost.com | Work at washingtonpost.com
The Washington Post: Subscribe | Subscriber Services | Advertise | Electronic Edition | Online Photo Store | The Washington Post Store | About The Post
The Washington Post Company: Information and Other Post Co. Websites