THE WALL STREET JOURNAL.

Ethics Aside, Is NSA's Spy Tool Efficient?

By: Carl Bialik – June 15, 2013

Reports about the National Security Agency's program to collect vast amounts of data on personal electronic communications have created an uproar about the implications for privacy. But some statisticians and security experts have raised another objection: As a terror-fighting tool, it is highly inefficient and has some serious downsides.

Their reasoning: Any automated approach to spotting something rare necessarily produces false positives. That means for every correctly identified target, many more alarms that go off will prove to be incorrect. So if there are vastly more innocent people than would-be terrorists whose communications are monitored, even an extremely accurate test would ensnare many non-terrorists.

A Ph.D. candidate in computational ecology wrote on his blog last week that even a very accurate algorithm for identifying terrorist communications could produce about 10,000 false positives for every real "hit," creating a haystack of false leads to chase in order to find every needle. Several media reports repeated the figure, and some experts agreed. "The false positives will kill you in this kind of system," said Bruce Schneier, a security technologist at U.K. telecommunications company BT Group PLC.

Some statisticians, though, disagreed—while emphasizing that they were commenting on the question of whether the tool was effective and not on its implications for privacy and civil liberties. They also said their answer depends in part on a host of questions that either can't be answered, or that the NSA won't answer—for instance, how effective its algorithm is at distinguishing dangerous communications from innocuous ones. An NSA spokeswoman declined to comment beyond recent statements released by the agency and the office of the Director of National Intelligence calling its monitoring tools "effective" and "important."

The value of these monitoring tools also depends on another question, though, that can be addressed: whether sifting for terrorism could be similar to screening for a rare disease. Corey Chivers, the fourth-year graduate student at McGill University in Montreal who wrote the blog post noted above, based his calculation on assuming the two kinds of screening are identical mathematically. The thinking goes like this: If you apply a test for something very rare to a population, even if it can be relied on to identify people who possess certain traits, there are so many more who don't that the false positives will vastly outnumber the correct hits.

The details of Mr. Chivers's calculation were based on some assumptions he made—for instance, that just one in one million people whose communications are being monitored are terrorists. If the prevalence is higher—depending on how one defines a terrorist—there will be fewer false positive tests for every true positive test.

Mr. Chivers said he never intended his proposed ratio to be very precise. "I found it interesting how a few people have picked up on the number itself," he said in an interview. "It's amazing how that will get replicated out of context of it being a thought experiment rather than a hard estimate of a number." Still, he said, his point stands: "Even if those numbers were vastly different, you're still going to have a large false-positive rate."

Several statisticians agreed with his basic point. Even if the NSA's algorithm "is terribly clever and has a very high sensitivity and specificity, it cannot avoid having an immense false-positive rate," said Peter F. Thall, a biostatistician at the University of Texas' M.D. Anderson Cancer Center. In his arena, false positives mean patients may get tests or treatment they don't need. For the NSA, false positives could mean innocent people are monitored, detained, find themselves on no-fly lists or are otherwise inconvenienced, and that the agency spends resources inefficiently.

Politics Counts

Others, though, noted a key difference between terrorism and, say, a needle in a haystack: Terrorists tend to talk to each other in a way that needles don't. So by analyzing a network of communications, the NSA could be ferreting out clues from more than just the messages' particulars.

Some statisticians suggested the NSA likely also is combining other evidence and signals with monitored communications to refine its hunt. "There is some prior information about who is a terrorist or where he or she is coming from that they could use to increase their chances," said Christopher H. Schmid, a biostatistician at Brown University's Center for Evidence- Based Medicine. "You start out with little idea what is going on, but by gathering more information you can refine your diagnosis."Tom Wallace, an analyst at the Center for Advanced Defense Studies, a Washington, D.C., think tank, offered hypothetical numbers to illustrate how even a one-in-10,000 finding could be helpful. "An analyst's knowledge of the region or field being studied could immediately narrow it down to 5,000," he said. Imagery analysis and a tip from an ally could bring it down to, say, 500. "Obviously all these numbers are made up, but hopefully they illustrate how statistical analysis is one element of the tool kit, not a magic wand," Mr. Wallace said.

But Jim Harper, director of information-policy studies at the libertarian Cato Institute think tank, said the cost of false positives is prohibitive.

"If you're mining for potential buyers of purses and it costs a dollar to send a catalog, your algorithm might be reliable if its false-positive rate is 95%. If you're mining for potential terrorists, the expense of getting it wrong is much higher"—in wasted investigation time and expense, and in potential civil-liberties violations.

Mr. Schneier asked a cost-effectiveness question: "Is this the absolute best thing to do with that money?"

Prof. Thall flipped the question, pointing out that any algorithm hunting for terrorists would turn up some number of false positives -- probably a large one. As to whether that should rule out using algorithms, though, he says, "I would very much like to know what alternative they might suggest. With regard to identifying terrorist attacks originating in the U.S.A. before they are carried out, there is no free lunch, and we simply can't have it both ways."