Mysterious communication connections by top 500 Android apps have no effect on user experience

November 19 2015, by Larry Hardesty

MIT researchers have shown that some phone apps engage in "covert communication." "There might be a very good reason for this covert communication. We are not trying to say that it has to be eliminated. We're just saying the user needs to be informed," says Julia Rubin, a postdoc in MIT's Computer Science and Artificial Intelligence Laboratory. Credit: Jose-Luis Olivares/MIT

MIT researchers have found that much of the data transferred to and from the 500 most popular free applications for Google Android cellphones make little or no difference to the user's experience.

Of those "covert" communications, roughly half appear to be initiated by standard Android analytics packages, which report statistics on usage patterns and program performance and are intended to help developers improve applications.

"The interesting part is that the other 50 percent cannot be attributed to analytics," says Julia Rubin, a postdoc in MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL), who led the new study. "There might be a very good reason for this covert communication. We are not trying to say that it has to be eliminated. We're just saying the user needs to be informed."

The researchers reported their findings last week at the IEEE/ACM International Conference on Automated Software Engineering. Joining Rubin on the paper are Martin Rinard, a professor of computer science and engineering at MIT; Michael Gordon, who received his PhD in electrical engineering and computer science in 2010 and remained at CSAIL as a researcher until last July; and Nguyen Nguyen of the engineering firm UWin Software.

Different operations performed by the same mobile app may require outside communication, and rather than try to coordinate shared access to a single communication channel, the app will typically open a separate communication channel for each operation.

The researchers analyzed the number of communication channels opened by the 500 most popular mobile apps and found that roughly 50 percent of them appear to have no bearing on the user experience. That doesn't necessarily translate directly to the quantity of data exchanged over those channels, but for short sessions of application use, the portion of transmitted data irrelevant to user experience is also as much as 50 percent.

Across longer sessions, in which large files are transferred to the phone—by, say, music- or video-streaming services—the percentage of data transmitted covertly steadily diminishes. But covert communication channels remain open.

Piercing the veil

Mobile applications are usually proprietary: Their source code is not publicly available, and their developers frequently take pains to disguise the details of the programs' execution, a technique known as obfuscation.

But all programs written for the Android operating system have to use a standard set of procedures when interacting with a phone's hardware. So it's possible to determine which portions of an Android app control what's displayed on the screen or piped through the speakers, and which portions open up communication channels.

By mapping out all of the possible ways that data can flow through an application, the researchers' analytic tools can determine whether a given command to open a communication channel will result in a control signal that goes to either the display or the speaker—and whether it won't.

To validate their analytic technique, the researchers produced modified versions of 47 of the top 100 Android apps, in which the communication channels that their tools identified as covert were disabled. They then conducted a series of usability studies in which pairs of subjects compared the performance of the modified and unmodified versions of the program.

In 30 of the 47 applications, subjects could detect no difference between the two versions. In nine instances, advertisements were missing but program execution was unaffected. And in three cases, the subjects characterized the differences as "minor." For instance, a popular free flashlight application offers users the option of buying functionality upgrades, and on the upgrade purchase screen, an icon was missing.

Five of the applications stopped working entirely. But in at least one of those cases, the problem was with protections that the developer had put in place to prevent reselling of its proprietary software. The reasons for the failure of the other four were not clear.

Who's listening?

The researchers also analyzed data traffic from a few of the more popular apps, gleaning some insight about the possible purposes of their covert communications. A Wal-Mart app, for instance, allows users to scan the bar codes of products on the shelves of Wal-Mart stores and retrieve their prices. But every time it does that, it also sends information to a server that appears to be associated with eBay. Disabling that connection had no effect on the app's behavior.

Interestingly, Candy Crush Saga, a game that got some bad press a few years ago for apparent privacy violations, was one of the very few apps that appeared to engage in no covert communication. "They've become a model citizen," Rubin says.

The analytic tools that the researchers used were based on an earlier project from Rinard's group, which Gordon had led. The U.S. Defense Advanced Research Projects Agency had sponsored research teams at nine universities in a four-year competition to develop techniques for identifying malicious software, or malware, in mobile applications developed for the Department of Defense by outside contractors. The DOD adopted the MIT team's system, which provided the basic framework for the new analysis.

"I think it's great work," says Omer Tripp, the technical lead on mobile security and privacy at IBM's T. J. Watson Research Center. "Where there's an element of surprise—and promise—is in the fact that you can't really localize all these covert channels to advertising and analytics, which is what one would intuitively expect."

Tripp speculates that some of the covert communication may be anticipatory, in an attempt to guard against interruptions in Internet connectivity. "You may imagine that the application may want to be more resilient and go on functioning without reporting a problem," he says. "Which, when you think about it, is an interesting opportunity for optimization. Perhaps some users say, 'If the app is willing to function without Internet connectivity, and I have a limited data plan, or I'm abroad and don't want to use the Internet, I want to know that it still knows how to do that.' The study sort of gives us the hope that in many cases this type of optimization could apply."

Tripp also points to a recent episode in which malicious hackers modified publicly distributed tools designed for iPhone application developers so that they injected malicious code into any application they were used to create. Even the best-intentioned developers could thus be party to breaches of security or privacy. "These are definitely use cases that are of interest to us at IBM," he says.

More information: Covert Communication in Mobile Applications. people.csail.mit.edu/mjulia/pu … pplications_2015.pdf

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

Citation: Mysterious communication connections by top 500 Android apps have no effect on user experience (2015, November 19) retrieved 4 May 2024 from https://phys.org/news/2015-11-mysterious-android-apps-effect-user.html