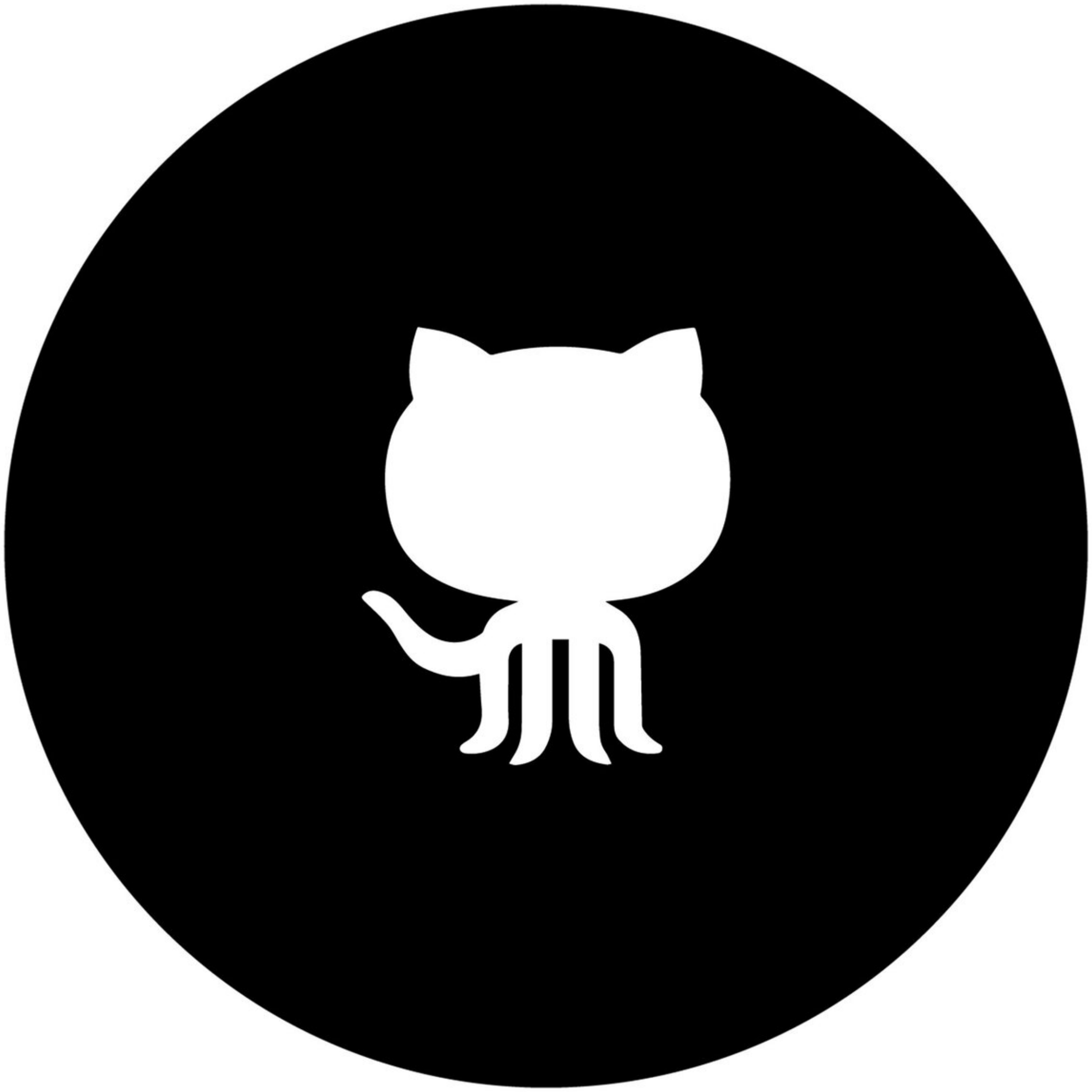
TechXplore

Software flaws often first reported on social media networks, researchers find

May 4 2020, by Allan Brettman



Credit: CC0 Public Domain

Software vulnerabilities are more likely to be discussed on social media before they're revealed on a government reporting site, a practice that could pose a national security threat, according to computer scientists at the U.S. Department of Energy's Pacific Northwest National Laboratory.

At the same time, those vulnerabilities present a cybersecurity opportunity for governments to more closely monitor social media discussions about software gaps, the researchers assert. Their findings were published recently in the journal *PLOS One*.

"Some of these software vulnerabilities have been targeted and exploited by adversaries of the United States. We wanted to see how discussions around these vulnerabilities evolved," said lead author Svitlana Volkova, senior research scientist in the Data Sciences and Analytics Group at PNNL. "Social cybersecurity is a huge threat. Being able to measure how different types of vulnerabilities spread across platforms is really needed."

Social media—especially GitHub—leads the way

Their research showed that a quarter of social media discussions about software vulnerabilities from 2015 through 2017 appeared on social media sites before landing in the National Vulnerability Database, the official U.S. repository for such information. Further, for this segment of vulnerabilities, it took an average of nearly 90 days for the gap discussed on social media to show up in the national database.

The research focused on three social platforms—GitHub, Twitter and Reddit—and evaluated how discussions about software vulnerabilities spread on each of them. The analysis showed that GitHub, a popular networking and development site for programmers, was by far the most likely of the three sites to be the starting point for discussion about software vulnerabilities.

It makes sense that GitHub would be the launching point for discussions about software vulnerabilities, the researchers wrote, because GitHub is a platform geared towards software development. The researchers found that for nearly 47 percent of the vulnerabilities, the discussions started on GitHub before moving to Twitter and Reddit. For about 16 percent of the vulnerabilities, these discussions started on GitHub even before they are published to official sites.

Codebase vulnerabilities are common

The research points at the scope of the issue, noting that nearly all commercial software codebases contain open-source sharing and that nearly 80 percent of codebases include at least one vulnerability. Further, each commercial software codebase contains an average of 64 vulnerabilities. The National Vulnerability Database, which curates and publicly releases vulnerabilities known as Common Vulnerabilities and Exposures "is drastically growing," the study says, "and includes more than 100,000 known vulnerabilities to date."

In their paper, the researchers discuss which U.S. adversaries might take note of such vulnerabilities. They mention Russia, China and others and noted that there are differences in usage of the three platforms within those countries when exploiting software vulnerabilities.

According to the study, cyberattacks in 2017 later linked to Russia involved more than 200,000 victims, affected more than 300,000 computers, and caused about $4 billion in damages.

"These attacks happened because there were known vulnerabilities present in modern software," the study says, "and some Advanced Persistent Threat groups effectively exploited them to execute a cyberattack."

Bots or human: Both pose a threat

The researchers also distinguished between social media traffic generated by humans and automated messages from bots. A social media message crafted by an actual person and not generated by a machine will likely be more effective at raising awareness of a software vulnerability, the researchers found, emphasizing that it was important to differentiate the two.

"We categorized users as likely bots or humans, by using the Botometer tool," the study says, "which uses a wide variety of user-based, friend, social network, temporal, and content-based features to perform bot vs. human classification."

The tool is especially useful in separating bots from human discussions on Twitter, a platform that the researchers noted can be helpful for accounts seeking to spread an agenda. Also regarding Twitter, the researchers found a subset of its users—for example FireEye, The Best Linux Blog In the Unixverse, The Hacker News and individual accounts belonging to cybersecurity experts—focused on news about software vulnerabilities.

Ultimately, awareness of social media's ability to spread information about software vulnerabilities provides a heads-up for institutions, the study says.

"Social media signals preceding official sources could potentially allow institutions to anticipate and prioritize which vulnerabilities to address first," it says. "Furthermore, quantification of the awareness of vulnerabilities and patches spreading in online social environments can provide an additional signal for institutions to utilize in their open source risk-reward decision making."

More information: Prasha Shrestha et al, Multiple social platforms reveal actionable signals for software vulnerability awareness: A study of GitHub, Twitter and Reddit, *PLOS ONE* (2020). DOI: 10.1371/journal.pone.0230250

Provided by Pacific Northwest National Laboratory

Citation: Software flaws often first reported on social media networks, researchers find (2020, May 4) retrieved 7 May 2024 from https://techxplore.com/news/2020-05-software-flaws-social-media-networks.html