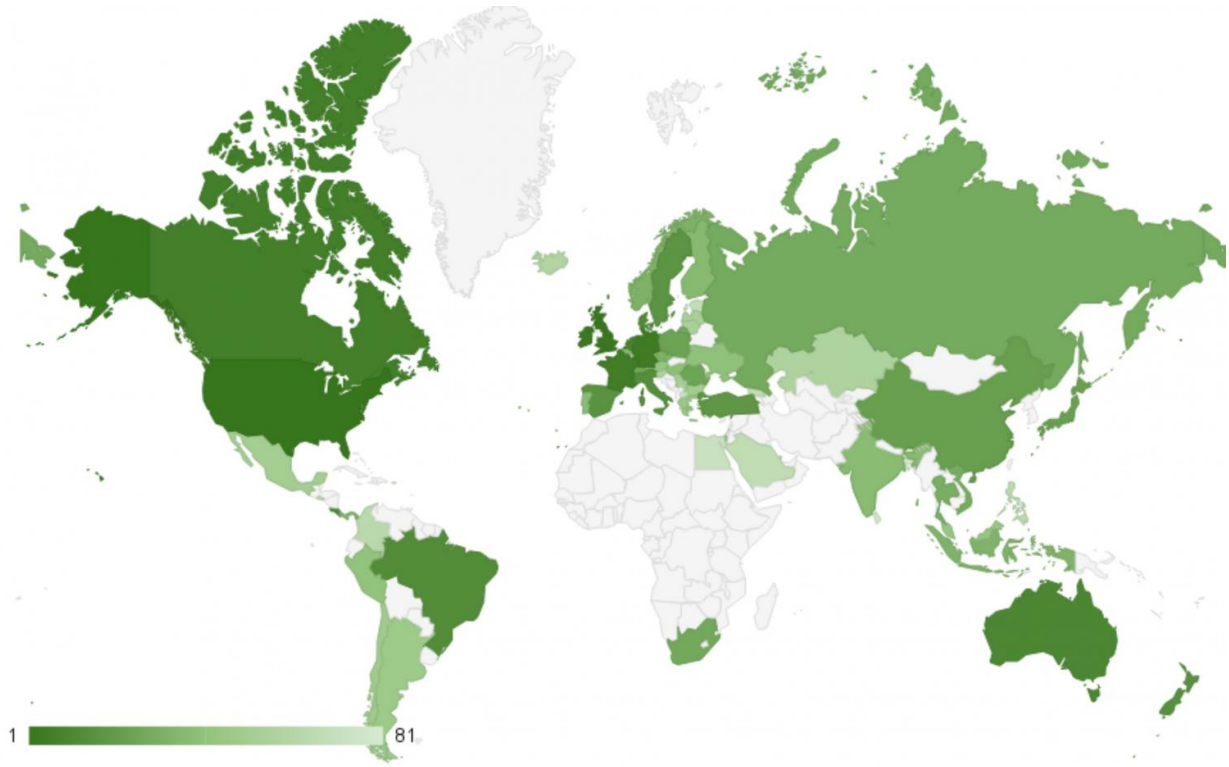
Study finds 'lurking malice' in cloud hosting services

October 18 2016



This map shows locations where the impacts of bad repositories (Bars) occur. Credit: Credit: Xiaojing Liao, Georgia Tech

A study of 20 major cloud hosting services has found that as many as 10 percent of the repositories hosted by them had been compromised - with several hundred of the "buckets" actively providing malware. Such bad content could be challenging to find, however, because it can be rapidly

assembled from stored components that individually may not appear to be malicious.

To identify the bad content, researchers created a scanning tool that looks for features unique to the bad repositories, known as "Bars." The features included certain types of redirection schemes and "gatekeeper" elements designed to protect the malware from scanners. Researchers from the Georgia Institute of Technology, Indiana University Bloomington and the University of California Santa Barbara conducted the study.

Believed to be the first systematic study of cloud-based malicious activity, the research will be presented October 24 at the ACM Conference on Computer and Communications Security in Vienna, Austria. The work was supported in part by the National Science Foundation.

"Bad actors have migrated to the cloud along with everybody else," said Raheem Beyah, a professor in Georgia Tech's School of Electrical and Computer Engineering. "The bad guys are using the cloud to deliver malware and other nefarious things while remaining undetected. The resources they use are compromised in a variety of ways, from traditional exploits to simply taking advantage of poor configurations."

Beyah and graduate student Xiaojing Liao found that the bad actors could hide their activities by keeping components of their malware in separate repositories that by themselves didn't trigger traditional scanners. Only when they were needed to launch an attack were the different parts of this malware assembled.

"Some exploits appear to be benign until they are assembled in a certain way," explained Beyah, who is the Motorola Foundation Professor and associate chair for strategic initiatives and innovation in the School of

Electrical and Computer Engineering. "When you scan the components in a piecemeal kind of way, you only see part of the malware, and the part you see may not be malicious."

In the cloud, malicious actors take advantage of how difficult it can be to scan so much storage. Operators of cloud hosting services may not have the resources to do the deep scans that may be necessary to find the Bars - and their monitoring of repositories may be limited by service-level agreements.

While splitting the malicious software up helped hide it, the strategy also created a technique for finding the "bad buckets" hosting it, Beyah said. Many of the bad actors had redundant repositories connected by specific kinds of redirection schemes that allowed attacks to continue if one bucket were lost. The bad buckets also usually had "gatekeepers" designed to keep scanners out of the repositories, and where webpages were served, they had simple structures that were easy to propagate.

"We observed that there is an inherent structure associated with how these attackers have set things up," he explained. "For instance, the bad guys all had bodyguards at the door. That's not normal for cloud storage, and we used that structure to detect them."

The researchers began by studying a small number of known bad repositories to understand how they were being used. Based on what they learned, they created "BarFinder," a scanner tool that automatically searches for and detects features common to the bad repositories.

Overall, the researchers scanned more than 140,000 sites on 20 cloud hosting sites and found about 700 active repositories for malicious content. In total, about 10 percent of cloud repositories the team studied had been compromised in some way. The researchers notified the cloud hosting companies of their findings before publication of the study.

"It's pervasive in the cloud," said Beyah. "We found problems in every last one of the hosting services we studied. We believe this is a significant problem for the cloud hosting industry."

In some cases, the bad actors simply opened an inexpensive account and began hosting their software. In other cases, the malicious content was hidden in the cloud-based domains of well-known brands. Intermingling the bad content with good content in the brand domains protected the malware from blacklisting of the domain.

Beyah and Liao saw a wide range of attacks in the cloud hosted repositories, ranging from phishing and common drive-by downloads to fake antivirus and computer update sites. "They can attack you directly from these buckets, or they can redirect you to other malicious buckets or a series of malicious buckets," he said. "It can be difficult to see where the code is redirecting you."

To protect cloud-based repositories from these attacks, Beyah recommends the usual defenses, including patching of systems and proper configuration settings.

Looking ahead, the researchers hope to make BarFinder available to a broader audience. That could include licensing the technology to a security company, or making it available as an open-source tool.

"Attackers are very clever, and as we secure things and make the cloud infrastructure more challenging for them to attack, they will move onto something else," he said. "In the meantime, every system that we can secure makes the internet just a little bit safer."

More information: Xiaojing Liao, et al., "Lurking Malice in the Cloud: Understanding and Detecting Cloud Repository as a Malicious Service," ACM Conference on Computer and Communications Security

(CCS).

Provided by Georgia Institute of Technology