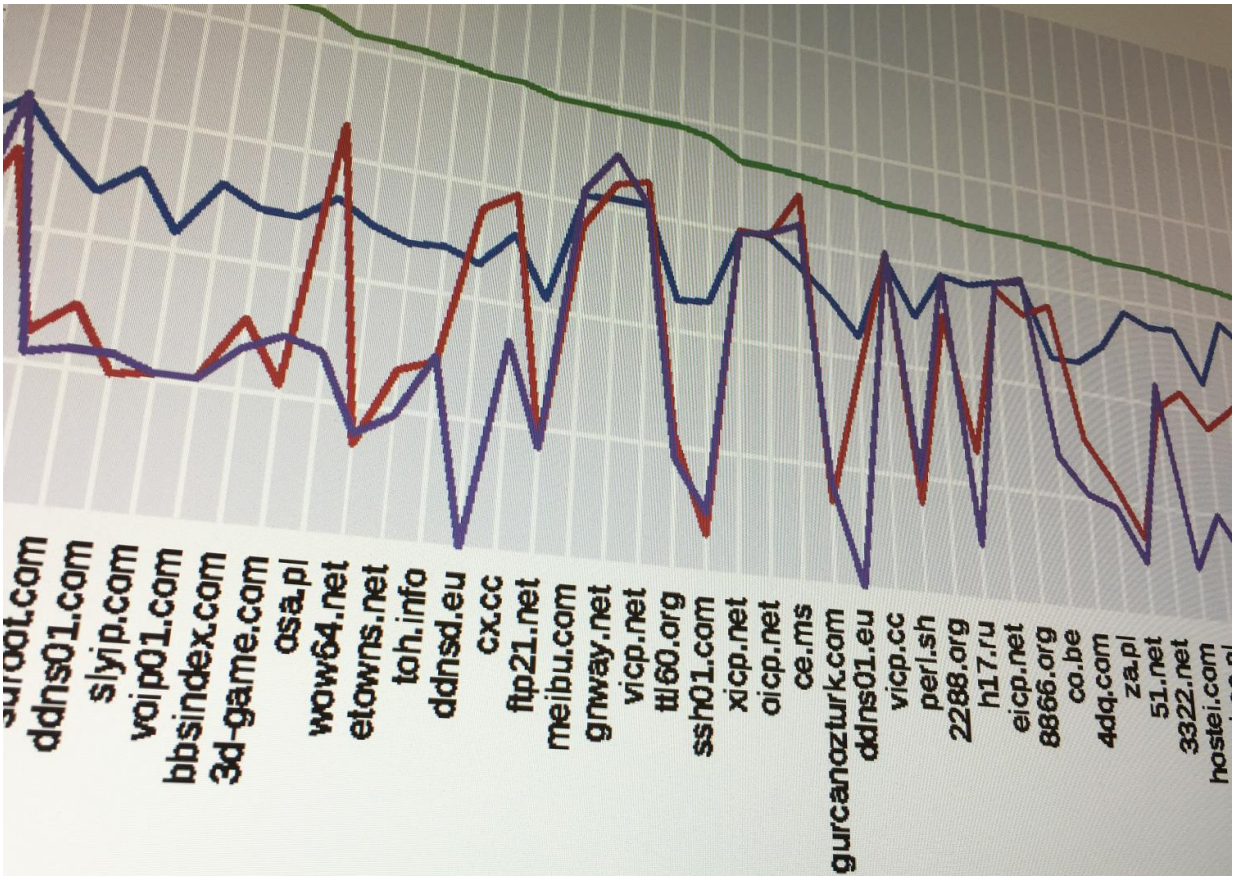
Network traffic provides early indication of malware infection

May 22 2017



Top domains queried by samples of malware studied by cybersecurity researchers at the Georgia Institute of Technology. Credit: Georgia Tech

By analyzing network traffic going to suspicious domains, security

administrators could detect malware infections weeks or even months before they're able to capture a sample of the invading malware, a new study suggests. The findings point toward the need for new malware-independent detection strategies that will give network defenders the ability to identify network security breaches in a more timely manner.

The strategy would take advantage of the fact that malware invaders need to communicate with their command and control computers, creating network traffic that can be detected and analyzed. Having an earlier warning of developing malware infections could enable quicker responses and potentially reduce the impact of attacks, the study's researchers say.

"Our study shows that by the time you find the malware, it's already too late because the network communications and domain names used by the malware were active weeks or even months before the actual malware was discovered," said Manos Antonakakis, an assistant professor in the School of Electrical and Computer Engineering at the Georgia Institute of Technology. "These findings show that we need to fundamentally change the way we think about network defense."

Traditional defenses depend on the detection of malware in a network. While analyzing malware samples can identify suspicious domains and help attribute network attacks to their sources, relying on samples to drive defensive actions gives malicious actors a critical time advantage to gather information and cause damage. "What we need to do is minimize the amount of time between the compromise and the detection event," Antonakakis added.

The research, which will be presented May 24 at the 38th IEEE Security and Privacy Symposium in San Jose, California, was supported by the U.S. Department of Commerce, the National Science Foundation, the Air Force Research Laboratory and the Defense Advanced Research

Projects Agency. The project was done in collaboration with EURECOM in France and the IMDEA Software Institute in Spain - whose work was supported by the regional government of Madrid and the government of Spain.

In the study, Antonakakis, Graduate Research Assistant Chaz Lever and colleagues analyzed more than five billion network events from nearly five years of network traffic carried by a major U.S. internet service provider (ISP). They also studied domain name server (DNS) requests made by nearly 27 million malware samples, and examined the timing for the re-registration of expired domains - which often provide the launch sites for malware attacks.

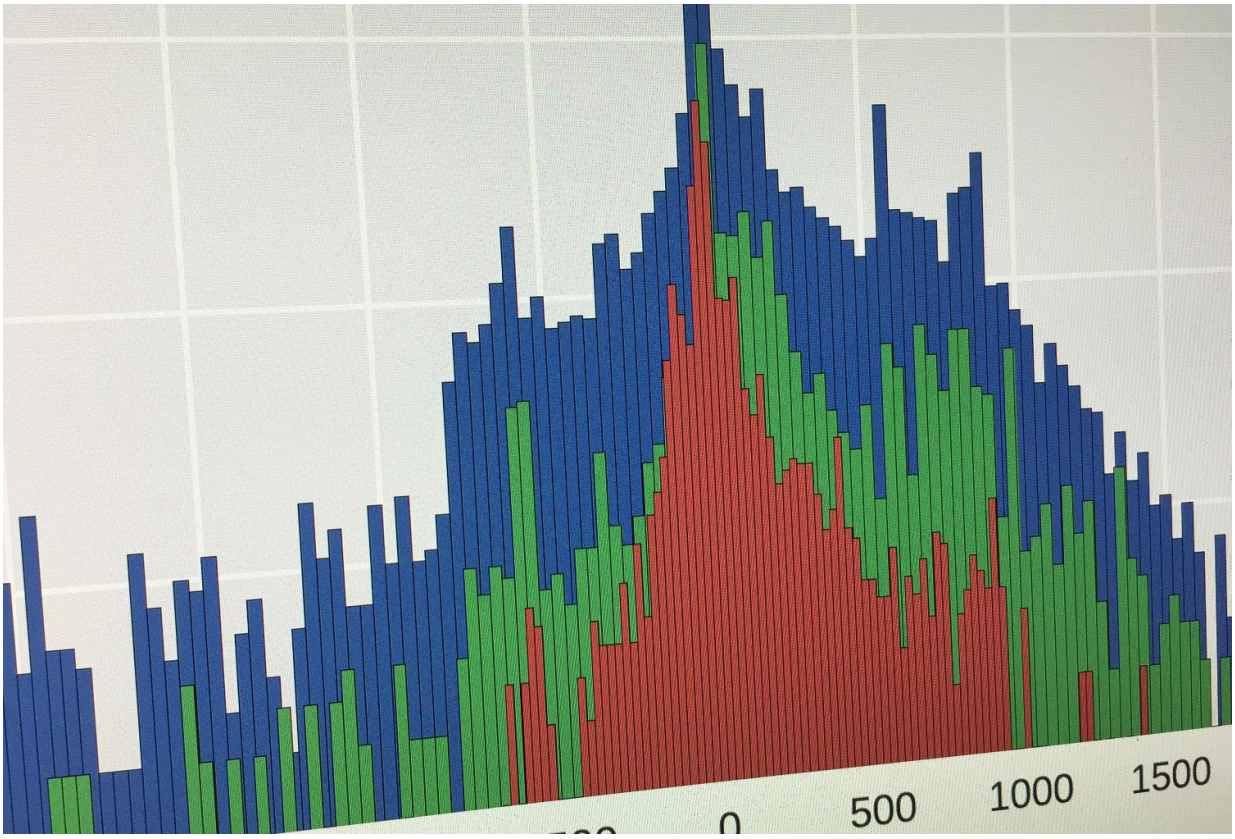


Chart shows the time difference between when malware signals were detected in

the network traffic of a major ISP and when the malware appeared on black lists. Credit: Georgia Tech

"There were certain networks that were more prone to abuse, so looking for traffic into those hot spot networks was potentially a good indicator of abuse underway," said Lever, the first author of the paper and a student in Georgia Tech's School of Electrical and Computer Engineering. "If you see a lot of DNS requests pointing to hot spots of abuse, that should raise concerns about potential infections."

The researchers also found that requests for dynamic DNS also related to bad activity, as these often correlate with services used by bad actors because they provide free domain registrations and the ability to add quickly add domains.

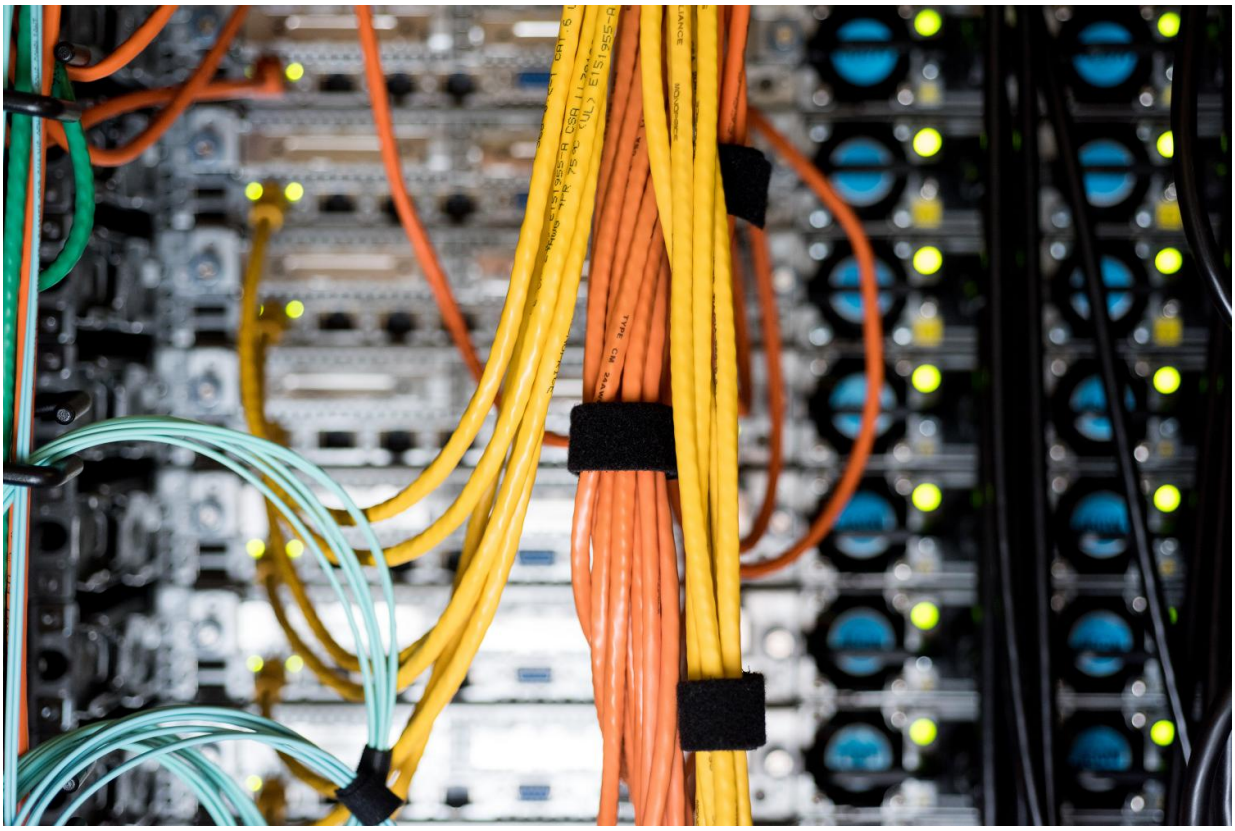
The researchers had hoped that the registration of previously expired domain names might provide a warning of impending attacks. But Lever found there was often a lag of months between when expired domains were re-registered and attacks from them began.

The research required development of a filtering system to separate benign network traffic from malicious traffic in the ISP data. The researchers also conducted what they believe is the largest malware classification effort to date to differentiate the malicious software from potentially unwanted programs (PUPs). To study similarities, they assigned the malware to specific "families."

By studying malware-related network traffic seen by the ISPs prior to detection of the malware, the researchers were able to determine that malware signals were present weeks and even months before new malicious software was found. Relating that to human health,

Antonakakis compares the network signals to the fever or general feeling of malaise that often precedes identification of the microorganism responsible for an infection.

"You know you are sick when you have a fever, before you know exactly what's causing it," he said. "The first thing the adversary does is set up a presence on the internet, and that first signal can indicate an infection. We should try to observe that symptom first on the network because if we wait to see the malware sample, we are almost certainly allowing a major infection to develop."



By analyzing network traffic going to suspicious domains, security administrators could detect malware infections weeks or even months before they're able to capture a sample of the invading malware, Georgia Tech researchers have found. Credit: Fitrah Hamid, Georgia Tech

In all, the researchers found more than 300,000 malware domains that were active for at least two weeks before the corresponding malware samples were identified and analyzed.

But as with human health, detecting a change indicating infection requires knowledge of the baseline activity, he said. Network administrators must have information about normal network traffic so they can detect the abnormalities that may signal a developing attack. While many aspects of an attack can be hidden, malware must always communicate back to those who sent it.

"If you have the ability to detect traffic in a network, regardless of how the malware may have gotten in, the action of communicating through the network will be observable," Antonakais said. "Network administrators should minimize the unknowns in their networks and classify their appropriate communications as much as possible so they can see the bad activity when it happens."

Antonakakis and Lever hope their study will lead to development of new strategies for defending computer networks.

"The choke point is the network traffic, and that's where this battle should be fought," said Antonakakis. "This study provides a fundamental observation of how the next generation of defense mechanisms should be designed. As more complicated attacks come into being, we will have to become smarter at detecting them earlier."

More information: Chaz Lever, et al., "A Lustrum of Malware Network Communication: Evolution and Insights," 38th IEEE Security and Privacy Symposium, 2017.

Provided by Georgia Institute of Technology