How artificial intelligence is taking on ransomware

28 June 2017, by Anick Jesdanun

In this Monday, May 15, 2017, file photo, employees watch electronic boards to monitor possible ransomware cyberattacks at the Korea Internet and Security Agency in Seoul, South Korea. Unable to rely on good human behavior, computer security experts are developing software techniques to fight ransomware. But getting these protections in the hands of users is challenging. (Yun Dong-jin/Yonhap via AP, File)

Twice in the space of six weeks, the world has suffered major attacks of ransomware—malicious software that locks up photos and other files stored on your computer, then demands money to release them.

It's clear that the world needs better defenses, and fortunately those are starting to emerge, if slowly and in patchwork fashion. When they arrive, we may have artificial intelligence to thank.

Ransomware isn't necessary trickier or more dangerous than other malware that sneaks onto your computer, but it can be much more aggravating, and at times devastating. Most such infections don't get in your face about taking your digital stuff away from you the way ransomware does, nor do they shake you down for hundreds of dollars or more.

Despite those risks, many people just aren't good at keeping up with security software updates. Both recent ransomware attacks walloped those who failed to install a Windows update released a few months earlier.

Watchdog security software has its problems, too. With this week's ransomware attack , only two of about 60 security services tested caught it at first, according to security researchers.

"A lot of normal applications, especially on Windows, behave like malware, and it's hard to tell them apart," said Ryan Kalember, an expert at the California security vendor Proofpoint.

HOW TO FIND MALWARE

In the early days, identifying malicious programs such as viruses involved matching their code against a database of known malware. But this technique was only as good as the database; new malware variants could easily slip through.

So security companies started characterizing malware by its behavior. In the case of ransomware, software could look for repeated attempts to lock files by encrypting them. But that can flag ordinary computer behavior such as file compression.

Newer techniques involve looking for combinations of behaviors. For instance, a program that starts encrypting files without showing a progress bar on the screen could be flagged for surreptitious activity, said Fabian Wosar, chief technology officer at the New Zealand security company Emsisoft. But that also risks identifying harmful software too late, after some files have already been locked up.

An even better approach identifies malware using observable characteristics usually associated with

malicious intent—for instance, by quarantining a program disguised with a PDF icon to hide its true nature.

This sort of malware profiling wouldn't rely on exact code matches, so it couldn't be easily evaded. And such checks could be made well before potentially dangerous programs start running.

MACHINE VS. MACHINE

Still, two or three characteristics might not properly distinguish malware from legitimate software. But how about dozens? Or hundreds? Or even thousands?

For that, security researchers turn to machine learning, a form of artificial intelligence. The security system analyzes samples of good and bad software and figures out what combination of factors is likely to be present in malware.

As it encounters new software, the system calculates the probability that it's malware, and rejects those that score above a certain threshold. When something gets through, it's a matter of tweaking the calculations or adjusting the threshold. Now and then, researchers see a new behavior to teach the machine.

AN ARMS RACE

On the flip side, malware writers can obtain these security tools and tweak their code to see if they can evade detection. Some websites already offer to test software against leading security systems. Eventually, malware authors may start creating their own machine-learning models to defeat security-focused artificial intelligence.

Dmitri Alperovitch, co-founder and chief technology officer at the California vendor CrowdStrike, said that even if a particular system offers 99 percent protection, "it's just a math problem of how many times you have to deviate your attack to get that 1 percent."

Still, security companies employing machine learning have claimed success in blocking most malware, not just ransomware. SentinelOne even offers a $1 million guarantee against ransomware; it hasn't had to pay it yet.

A FUNDAMENTAL CHALLENGE

So why was ransomware still able to spread in recent weeks?

Garden-variety anti-virus software—even some of the free versions—can help block new forms of malware, as many are also incorporating behavioral-detection and machine-learning techniques. But such software still relies on malware databases that users aren't typically good at keeping up to date.

Next-generation services such as CrowdStrike, SentinelOne and Cylance tend to ditch databases completely in favor of machine learning.

But these services focus on corporate customers, charging $40 to $50 a year per computer. Smaller businesses often don't have the budget—or the focus on security—for that kind of protection.

And forget consumers; these security companies aren't selling to them yet. Though Cylance plans to release a consumer version in July, it says it'll be a tough sell—at least until someone gets attacked personally or knows a friend or family member who has.

As Cylance CEO Stuart McClure puts it: "When you haven't been hit with a tornado, why would you get tornado insurance?"