ARTICLE

DARPA Proves Automated Systems Can Detect, Patch Software Flaws at Machine Speed

Written by: Debbie Walkowski, David Holmes, John Hall

Date: October 23, 2016

According to the Defense Advanced Research Projects Agency (DARPA), it takes an average of 312 days for security pros to discover software vulnerabilities such as viruses, malware, and other attacks. In hacker time, that's a virtual eternity in which bad actors can wreak havoc within infected systems and steal information, all without being noticed.

DARPA would like to see that 312 days reduced to a matter of weeks, days—even seconds. How? By putting automated machines to the task.

That was the idea behind the DARPA Cyber Grand Challenge (CGC), a first-of-its-kind cyber capture the flag (CTF) competition in which the competitors were machines, not humans. Spanning two years, the project began in 2014 with 100 teams attempting to program computers to play the game. The challenge culminated in August 2016 in a final showdown among the top seven finalist teams. Individual, fully autonomous "cyber reasoning systems" competed, each composed of 2560 CPU cores and 16TB of RAM, along with automated fuzzing, symbolic execution, analysis, and management software.

DARPA's objective? To improve the state of cyber security "…by developing automated, scalable systems able to find and fix software vulnerabilities at machine speed," DARPA director Arati Prabhakar was quoted as saying in a DARPA press release[1]. The current process for finding

exploitable vulnerabilities and bugs in software is not automated. Security professionals spend thousands of hours searching millions of lines of code to find and patch software flaws. “Our goal is to break past the reactive patch cycle we’re living in today,” Prabhakar added.

For two decades, security pros have been honing their bug-hunting skills in CTF competitions at network security and hacking conferences. Teams of human players face off in a head-to-head race to discover, diagnose, and fix software flaws in real time. Each player controls a server (host) running an identical copy of unexplored code. During the game, players are given small, original programs called “challenge binaries” that contain vulnerabilities and flags to be protected or captured. Players must protect their digital flags by patching their own server software, keeping it healthy and functional, and scanning for and attacking opponents’ vulnerabilities to capture their flags. Players earn points for defending their server code and keeping their flags safe, keeping their software available and functioning normally, and capturing opponents’ flags; they lose points for damaging their own software and for losing flags.

Cyber CTF is a game of strategy and tactics that requires analytic skills, speed, and perseverance. When players find a vulnerability, they must decide what to do: Patch immediately? Don’t patch and watch the network? Scan opponents first? Tell no one? Build an obfuscated defense? In the real world today, it’s human players, not machines, who wrestle with and make these decisions.

The implications of automated machines someday being able to handle these decisions are enormous. The possibility of dramatically reducing the time it takes to find and fix software vulnerabilities is critical in a world where every conceivable object is being connected to the Internet. That includes everyday conveniences such as household appliances and cars as well as critical systems such as power grids, traffic lights, water supplies, and air traffic control systems.

In the mid-2000s, it was estimated that the world was running a trillion lines of code. The Internet of Things means a vast increase in the number and types of devices becoming “connected.” New code for these devices is being written every day—quite often by engineers who have little experience writing secure code to operate on the hostile Internet. That means vulnerabilities and attacks have the potential to expand at an alarming rate as more and more connected devices are produced. We are already seeing the results of this with DDoS attacks launched by IoT devices reaching above 600 Gbps at their target and much larger at their source.

So, what were the results of the CGC?

The DARPA CGC finals demonstrated that automated binary vulnerability analysis is technically possible, now; and as with other computational problems, it will only get faster and more capable with time. Each of the competitors blended automated fuzzing, with a library of well researched attack patterns, and symbolic execution, with its ability to "solve for the crash," to find and fix vulnerabilities in previously unknown challenge binaries. In several cases, competitors found and exploited vulnerabilities within minutes! All of the competitors delivered many variations of each attack as well. In some cases, they found variations of attacks for existing vulnerabilities which continued to work when the published fixes for those vulnerabilities were applied! On the other hand, a large number of known vulnerabilities in the challenge binaries were not found by any of the competitors. So none of these cyber reasoning systems yet provides a "general solution" to the problem of finding and exploiting vulnerabilities.

Interestingly, while each team used a similar high-level approach to finding, fixing, and exploiting vulnerabilities, their cyber reasoning systems had different strengths and weaknesses as shown in the competition results. Some were better at attacking, some better at fixing issues, some found the most vulnerabilities, and some simply played the game better.

These results imply that there is much fertile ground available for research and for tuning or extending the various parts of these complex systems. And that network defenders and product manufacturers are fated for interesting times ahead.

The future possibility of this automation is fun to think about, and will be even more exciting to witness when it does happen. But, the world has a lot of catching up to do before this will be a reality. Automation is every tech shop's Holy Grail, more often than not stalled by "tech debt" that causes the application to break when things are automated. In a world where most applications are a conglomeration of old and new code bases, sometimes multiple languages, customized third-party plugins, out of date OSs, and lots of internal network dependencies, we might need to start fresh to make this a reality.

About F5 Labs

F5 Labs combines the expertise of our security researchers with the threat intelligence data we collect to provide actionable, global intelligence on current cyber threats—and to identify future trends. We look at everything from threat actors, to the nature and source of attacks, to post-attack analysis of significant incidents to create a comprehensive view of the threat landscape. From the newest malware variants to zero-day exploits and attack trends, F5 Labs is where you'll find the latest insights from F5's threat intelligence team.

[1] http://www.darpa.mil/news-events/2016-08-05a

US Headquarters: 401 Elliott Ave W, Seattle, WA 98119 | 888-882-4447 // Americas: info@f5.com // Asia-Pacific: apacinfo@f5.com // Europe/Middle East/Africa: emeainfo@f5.com // Japan: f5j-info@f5.com