Cybersecurity and Deep Learning Applications

Berk Sunar

Worcester Polytechnic Institute

Machine Learning in Cybersecurity

- Intrusion Detection
- Phishing and Spam Email detection
- Malware detection in digital documents
- Click fraud detection
- Captcha and password guessing
- Fake account and Bot detection
- Fake news and Deepfakes generation and detection
- Adversarial Networks
- Side-channel Analysis
- …

Early Machine Learning Work

- 2007
  - Designed first technology to detect tampering and counterfeiting in chips manufactured overseas
  - First use of ML (PCA) in hardware security
  - Won IBM Pat Goldberg Best Paper Award in 2007
  - 680 academic citations

IEEE Spectrum 2008 Article

The Hunt for the Kill Switch

Are chip makers building electronic trapdoors in key military hardware? The Pentagon is making its biggest effort yet to find out

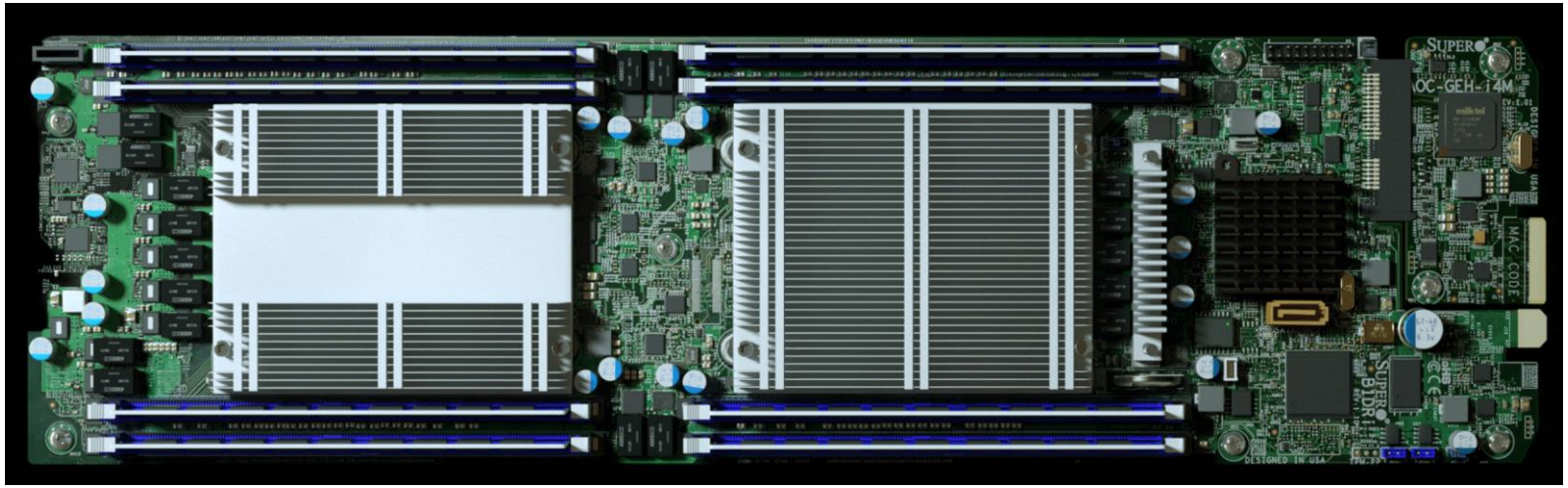
Posted 1 May 2008 | 19:57 GMT
By **SALLY ADEE**

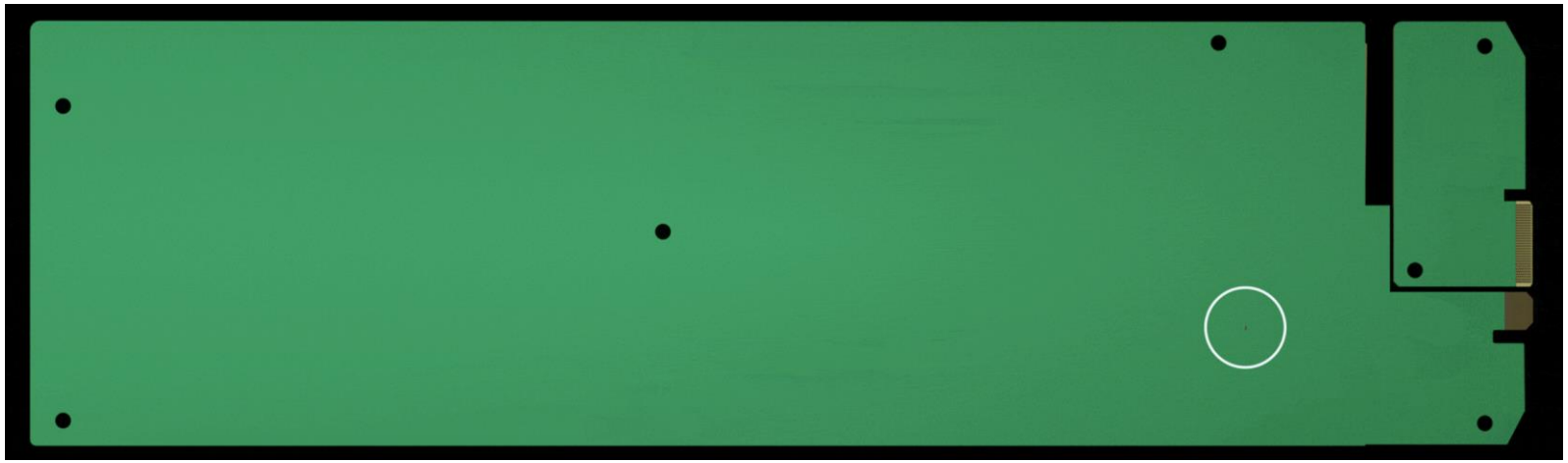
Trojan Detection Using IC Fingerprinting [IEEE Xplore]

Authors: Dakshi Agrawal (IBM Research - Watson), Pankaj Rohatgi (IBM Research - Watson), Selçuk Baktir (WPI), Deniz Karakoyunlu (WPI), Berk Sunar (WPI)

http://researcher.watson.ibm.com/researcher/view_group.php?id=5855

Hardware Trojans





October 4,2018

Amazon AWS Attack

- 2015:
  - First RSA Key Recovery Attack Performed on the Real AWS Amazon Cloud
  - Cross-VM: Key stolen from co-located VM OS through performance leakages
  - Extremely noisy environment
  - Extensive data processing and classification work to recover exact keys
  - Worked with Amazon AWS team to address the issue
  - Tons of media coverage
  - Similar work later on Microsoft Azure

AWS Attack Coverage

NETWORKWORLD Most read:

Home > Cloud Computing > Cloud Security

AWS RE:INVENT 2015

Researchers steal secret RSA encryption keys in Amazon's cloud

Credit: Thinkstock

RELATED

Is BYOK the key to secure cloud computing?

Review: Container wars: Rocket vs. Odin vs. Docker

Amazon CTO Vogels on competing with Microsoft in hybrid cloud, keeping...

on IDG Answers
How is data stored and accessed in the cloud?

Now-patched attack raises questions about security of cloud environments

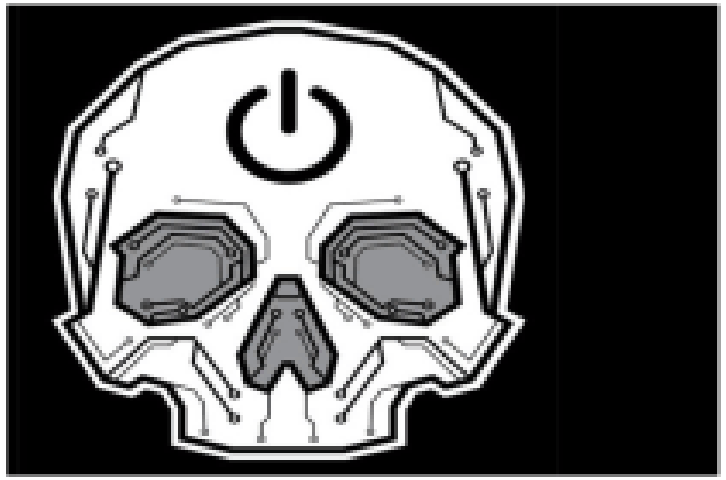
By Brandon Butler | Follow
Network World | Oct 6, 2015 8:40 AM PT

The Register

Security

Amazon boards windows against leet key-stealing neighbours

Want security? No no no don't go co-lo.



Amazon has patched a vulnerability that could have let users to steal the RSA keys of other co-located customers.

The complex attack - getting to CPU code cache isn't trivial - would, if successful, give an attacker a whole 2048-bit key used in other Elastic Compute Cloud instances.

More Recent Work

- Heavy shift of using machine learning/AI:
  - MasCat:
    - Scanning tool to detect malicious leakages from smartphone apps
    - First tool to scan app-store for stealthy leakage attacks, e.g. Rowhammer.
  - PerfWeb:
    - can tell which website you are visiting!
    - Breaks through Tor browser isolation
    - Heavy use of ML/AI models for improved classification
    - Media coverage in Vice News

Security

Boffins bag side-channel bugs before they bite

How to spot a side order of Rowhammer in a benign binary

By Richard Chirgwin 4 Jan 2017 at 06:04 SHARE ▼

Rowhammer and similar side-channel attacks aren't caught by anti-virus, so a bunch of US boffins have set about working out how to catch their signatures.

Once considered the stuff of laboratories and spies, side-channel attacks have become increasingly practical. Rowhammer, for example, is a software-only way to flip bits in one row of RAM by rapidly writing and re-writing bits in another row. Ultimately, it lets the attacker crash a kernel process to get root access.

The trio from America's Worcester Polytechnic Institute (WPI) in Massachusetts – Gorka Irazoqui, Thomas Eisenbarth, and Berk Sunar – have published a paper at the International Association for Cryptologic Research (IACR), presenting what they say is analogous to a virus-scan for side-channel attacks.

MasCat Press Coverage

MOTHERBOARD

PRIVACY

This Spy App Can See If You've Visited Whistleblowing Sites on the Dark Web

JORDAN PEARSON
May 18 2017, 11:00am

It's not in hackers' hands—yet.

SHARE TWEET

To stay off the radar when leaking information to the press, whistleblowers often turn to the dark web to mask their identity. But that's no match for a new malicious app that spies on your computer hardware, and can tell when you've visited whistleblower sites through the Tor Browser.

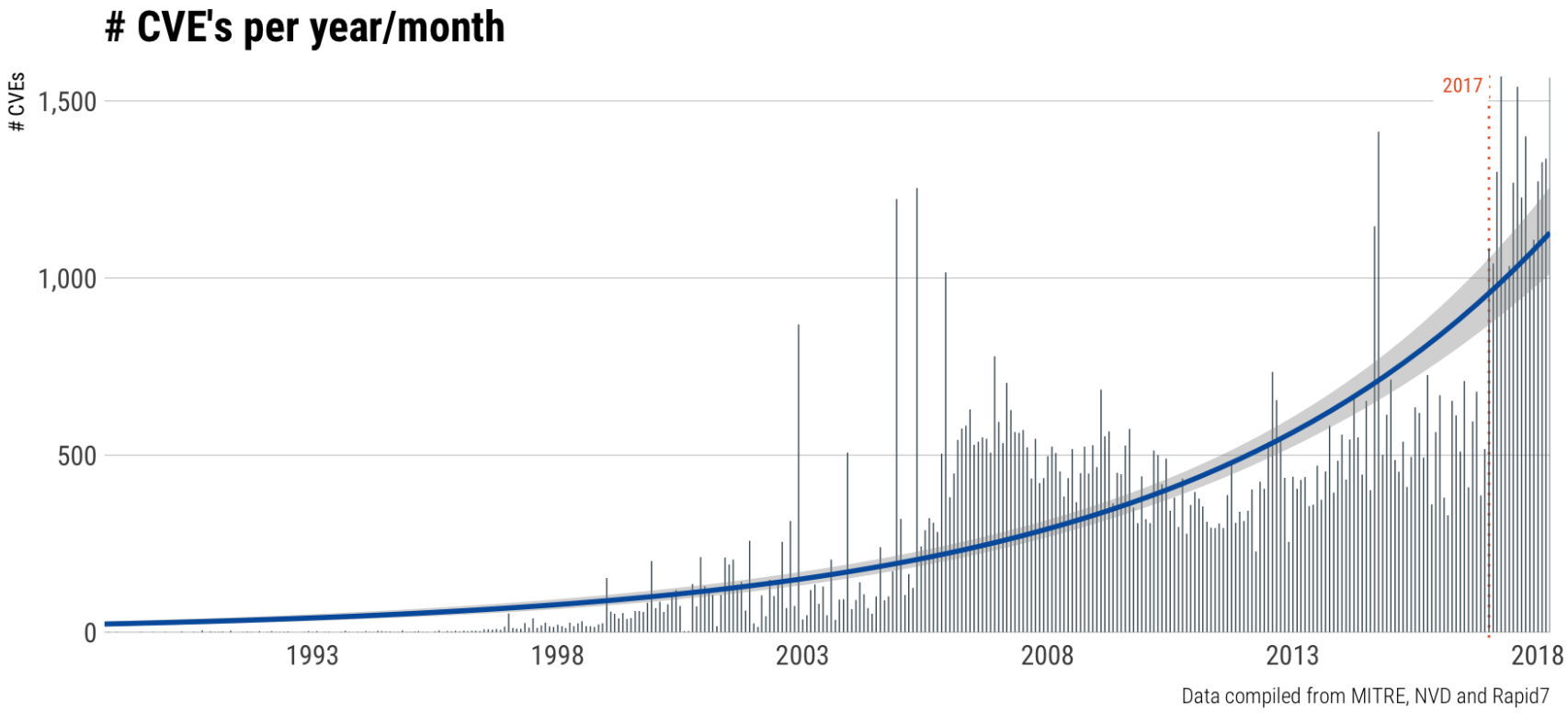
Thankfully, this revelation doesn't come from hackers. Instead, the app was developed by computer scientists at the Worcester Polytechnic Institute (WPI), and they uploaded a paper outlining their work to the arXiv preprint server last week. Their app makes use of a well-known attack in academic circles: if you carefully track and analyze the patterns of use on a computer's processor, you can piece together what the user is actually doing.

Now, the researchers have shown that it can be done with a malicious app running in the background on someone's machine, and a bit of AI.

"You might protect your browsing habits by going into incognito mode or using the Tor Browser—the traffic there is hidden from, say, your IT admin," said Berk Sunar, one of the study's co-authors, over the phone. "What we're showing here is that in that unprotected corporate environment, even using tools like Tor, your browsing history can be leaked in part to a monitoring authority."

PerfWeb Press Coverage in Vice News

Manual Vulnerability Analysis Does not Scale!



Source rapid7

DARPA CGC

- August 2016
- world's first all-machine cyber hacking tournament
- 100 teams consisting of some of the top security researchers
- $2 million, $1 million, and $750K awarded
- Capture the Flag testbed laden with hidden bugs
- Challenge: Find and patch vulnerable code in seconds!

DARPA DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

EXPLORE BY TAG

ABOUT US / OUR RESEARCH / NEWS / EVENTS / WORK WITH US /

Defense Advanced Research Projects Agency › Program Information

Cyber Grand Challenge (CGC)

Mr. Dustin Fraze

The need for automated, scalable, machine-speed vulnerability detection and patching is large and growing fast as more and more systems—from household appliances to major military platforms—get connected to and become dependent upon the internet. Today, the process of finding and countering bugs, hacks, and other cyber infection vectors is still effectively artisanal. Professional bug hunters, security coders, and other security pros work tremendous hours, searching millions of lines of code to find and fix vulnerabilities that could be taken advantage of by users with ulterior motives.

To help overcome these challenges, DARPA launched the Cyber Grand Challenge, a competition to create automatic defensive systems capable of reasoning about flaws, formulating patches and deploying them on a network in real time. By acting at machine speed and scale, these technologies may someday overturn today's attacker-dominated status quo. Realizing this vision requires breakthrough approaches in a variety of disciplines, including applied computer security, program analysis, and data visualization. Anticipated future benefits include:

- Expert-level software security analysis and remediation, at machine speeds on enterprise scales
- Establishment of a lasting R&D community for automated cyber defense
- Creation of a public, high-fidelity recording of real-time competition between automated cyber defense systems

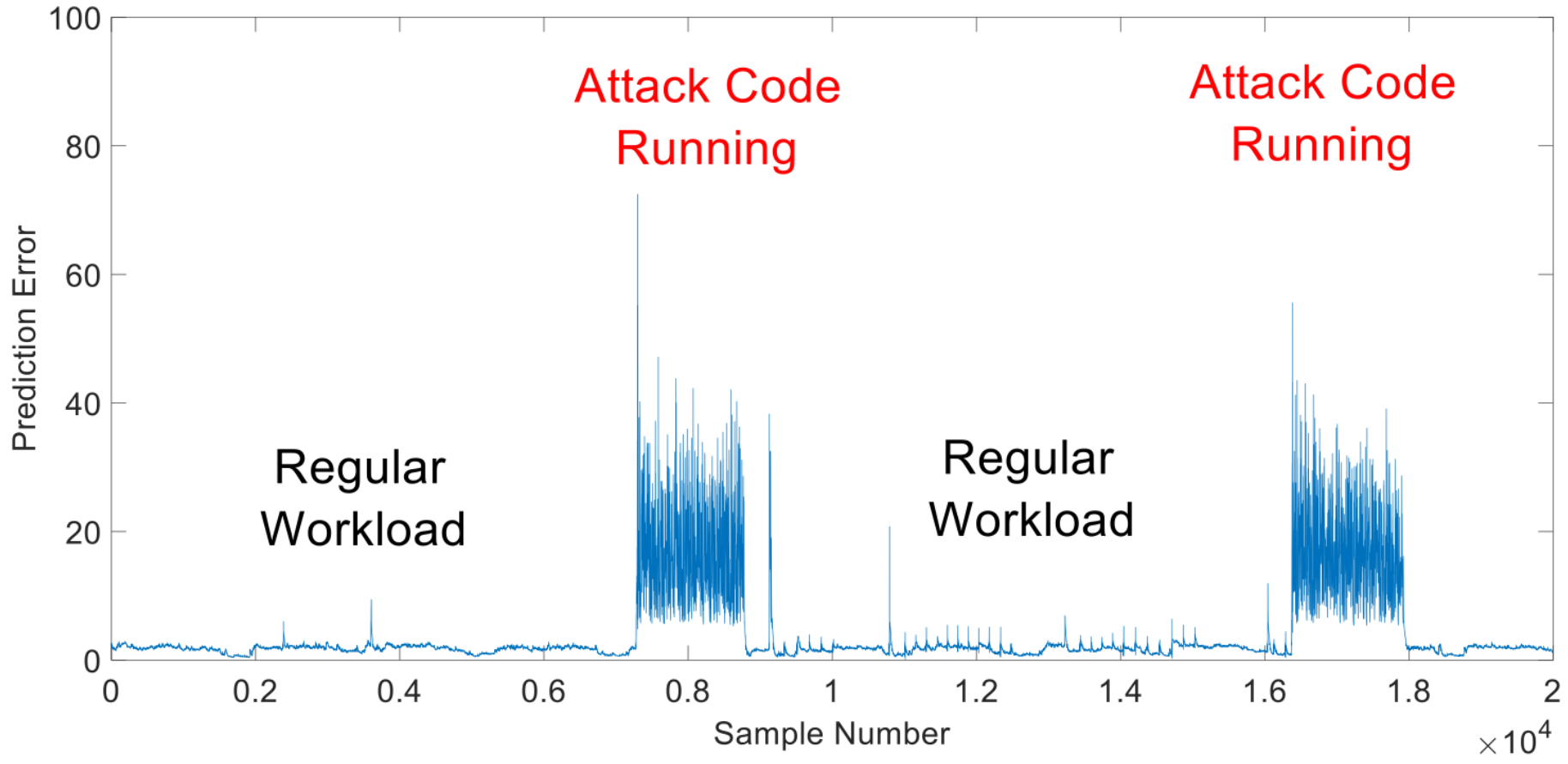
DARPA hosted the Cyber Grand Challenge Final Event—the world's first all-machine cyber hacking tournament—on August 4, 2016 in Las Vegas. Starting with over 100 teams consisting of some of the top security researchers and hackers in the world, DARPA pit seven teams against each other during the final event. During the competition, each team's Cyber Reasoning System (CRS) automatically identified software flaws, and scanned a purpose-built, air-gapped network to identify affected hosts. For nearly twelve hours, teams were scored based on how capably their systems protected hosts, scanned the network for vulnerabilities, and maintained the correct function of software. Prizes of $2 million, $1 million, and $750 thousand were awarded to the top three finishers.

FortuneTeller

- Aim: Detecting previously unknown micro-architectural attacks such as cache based side-channel attacks, e.g. Meltdown, Spectre and Rowhammer
- Anomaly Detection with RNN
- Unsupervised training LSTM with Phoronix Test Suite tests data
- 85% success rate to detect the attack on the server 10% false positive rate

Prediction with LSTM

Sample Attack



Split View in Applied Crypto Community

- Attacks require high level of intelligence
- Humans are intrinsically better in discovering new vulnerabilities
- ML may be used to automate existing attacks

- Deep Learning is ready for prime-time
- Attackers are already using DL
- Only way to scale new vulnerability discovery is to use Deep Learning

Teşekkürler!