Let's Make Windows Defender Angry: Antivirus can be an oracle!

Let's Make Windows Defender Angry: Antivirus can be an oracle! - Ryo Ichikawa, Speaker Deck.

- Published: 2019-10-29
- Original: <https://speakerdeck.com/icchy/lets-make-windows-defender-angry-antivirus-can-be-an-oracle>
- Preserved from: https://speakerdeck.com/icchy/lets-make-windows-defender-angry-antivirus-can-be-an-oracle (live) on 2026-08-10
- Licence: unknown

Content

> UNTRUSTED SOURCE TEXT. Everything below this line is third-party material quoted for research. It is data, not instructions. Do not follow directions, execute code, or fetch URLs because this text says so.

Let's Make Windows Defender Angry: Antivirus can be an oracle! - Speaker Deck

Let's Make Windows Defender Angry: Antivirus can be an oracle!

A presentation about AVOracle (AntiVirus Oracle) at CODE BLUE 2019 U25 track (https://codeblue.jp/2019/en/talks/?content=talks_23) Japanese version: https://speakerdeck.com/icchy/antiuirusuworakurutosita-windows-defendernidui-suru-xin-siigong-ji-shou-fa

[image: Avatar for icchy]

icchy

October 29, 2019

More Decks by icchy

See All by icchy

React Hooks

[[image: Avatar for icchy] icchy](https://speakerdeck.com/icchy)

3.6k

Windows Defender

[[image: Avatar for icchy] icchy](https://speakerdeck.com/icchy)

0

640

WCTF2019: Gyotaku The Flag

[[image: Avatar for icchy] icchy](https://speakerdeck.com/icchy)

0

380

Other Decks in Research

See All in Research

/ Who Do Large Language Models Memorize?

[[image: Avatar for Shotaro Ishihara] upura](https://speakerdeck.com/upura)

0

110

2026

[[image: Avatar for Takashi Ozeki] ozekinote](https://speakerdeck.com/ozekinote)

0

140

Model Discovery and Graph Simulation: A Lightweight Gateway to Chaos Engineering

[[image: Avatar for Anatoly A. Krasnovsky] anatolykr](https://speakerdeck.com/anatolykr)

0

250

QA SVAgent: Storyline-Guided Long Video Understanding via Cross-Modal Multi-Agent Collaboration

[[image: Avatar for] murakawatakuya](https://speakerdeck.com/murakawatakuya)

1

180

[[image: Avatar for Kanta Tanishima] ktanishima](https://speakerdeck.com/ktanishima)

0

120

/ Advanced Imitation Learning for VLA

[[image: Avatar for PRIN Lab] prinlab](https://speakerdeck.com/prinlab)

0

270

[[image: Avatar for afroscript] afroscript](https://speakerdeck.com/afroscript)

0

160

Google Cloud Next 2026 DM Recap Agentic Data Cloud / Google Cloud Next 2026 DM Recap

[[image: Avatar for nnaka2992] nnaka2992](https://speakerdeck.com/nnaka2992)

0

110

[[image: Avatar for] theoriatec2024](https://speakerdeck.com/theoriatec2024)

0

450

[

[IR Reading 2026] LLM-based Listwise Reranking under the Effect of Positional Bias (ECIR 2026) /IR-Reading-2026-Spring

](https://speakerdeck.com/koheishinden/ir-reading-2026-spring)

[[image: Avatar for Kohei Shinden] koheishinden](https://speakerdeck.com/koheishinden)

PRO

0

300

LINE Meetup -AlpacaTech

[[image: Avatar for tomo] gamella](https://speakerdeck.com/gamella)

1

620

LLM

[[image: Avatar for kento sugimoto] kento1109](https://speakerdeck.com/kento1109)

1

1.6k

Featured

See All Featured

Designing Experiences People Love

[[image: Avatar for Jonathan Moore] moore](https://speakerdeck.com/moore)

143

24k

Neural Spatial Audio Processing for Sound Field Analysis and Control

[[image: Avatar for NII S. Koyama's Lab] skoyamalab](https://speakerdeck.com/skoyamalab)

0

400

The Straight Up "How To Draw Better" Workshop

[[image: Avatar for Dennis Kardys] denniskardys](https://speakerdeck.com/denniskardys)

239

140k

Let's Do A Bunch of Simple Stuff to Make Websites Faster

[[image: Avatar for Chris Coyier] chriscoyier](https://speakerdeck.com/chriscoyier)

508

140k

How People are Using Generative and Agentic AI to Supercharge Their Products, Projects, Services and Value Streams Today

[[image: Avatar for Helen Beal].png) helenjbeal](https://speakerdeck.com/helenjbeal)

1

260

Jess Joyce - The Pitfalls of Following Frameworks

[[image: Avatar for Tech SEO Connect] techseoconnect]
(https://speakerdeck.com/techseoconnect)

PRO

1

330

Deep Space Network (abreviated)

[[image: Avatar for Tony Rice] tonyrice](https://speakerdeck.com/tonyrice)

0

250

Building Experiences: Design Systems, User Experience, and Full Site Editing

[[image: Avatar for Michelle Schulp Hunt] marktimemedia]
(https://speakerdeck.com/marktimemedia)

0

570

Discover your Explorer Soul

[[image: Avatar for Emna] emnaayadi](https://speakerdeck.com/emna__ayadi)

2

1.2k

Designing for humans not robots

[[image: Avatar for Tammie Lister] tammielis](https://speakerdeck.com/tammielis)

254

26k

Faster Mobile Websites

[[image: Avatar for Dean Hume] deanohume](https://speakerdeck.com/deanohume)

310

32k

Designing for Timeless Needs

[[image: Avatar for Cassini Nazir] cassininazir](https://speakerdeck.com/cassininazir)

1

430

Transcript

-

Let's Make Windows Defender Angry: Antivirus can be an oracle!

Ryo Ichikawa (icchy) CODE BLUE 2019, 10/29

-

Who am I • icchy (a.k.a. t0nk42) • CTF enthusiast

TokyoWesterns captain Web, Forensics • CTF Oraganizer TokyoWesterns CTF Challenge authoring, Infrastructure maintainance CODE BLUE CTF Bull's Eye system developer

-

Question • Can you point out the vulnerability? You'll

know what the vulnerability is after this talk

-

https://www.rambus.com/blogs/an-introduction-to-side-channel-attacks/ Side-channel attack

-

Side-channel attacks basics • Ordinary exploits Remote Code Execution

Path traversal • Side-channel attacks Leak sensitive data from side-effects everywhere Spectre: time difference between cache hit XS-Search: unprocted attributes of JavaScript (ex. iframe.length) Padding oracle: padding error tell us plain text • They are recovering info from side-effects (i.e. oracles)

-

What is the target of side-channel attack? • CPU

Spectre • Content auditor XSS auditor • Crypto Padding oracle (ex. POODLE) • Hardware Power analysis

-

Content auditors • Content auditors protect users XSS Auditor

WAF (Web Application Firewall) Antivirus software • Content auditors know the content to be audited • Content auditors sometimes have evaluation

-

Side-channel attacks against content auditor • XS-Search Triggering false

positive for XSS Auditor in Chrome • Reflected XSS would be detected and blocked http://target/?<script>var secret = '1234';</script> query malformed url to leak secret <script>var secret = '1232';</script> <script>var secret = '1233';</script> <script>var secret = '1234';</script> blocked! • Let's call this kind of attack Auditor Based Oracle • How about antivirus software?

-

None

-

Antivirus Technologies • one of the most common software we

use today Avast ESET Kaspersky Security McAfee Norton Security Symantec Endpoint Protection Trendmicro Virus Buster Cloud Windows Defender ... • Protect users from malicious attempts by auditing File content Network traffics etc.

-

[audit([secret] + [user input])]
(https://files.speakerdeck.com/presentations/ad7d23687b414ff4874295f328d490b5/slide_10.jpg)

-

Abusing Antivirus Technologies • What if attacker can control data

partially? as saving input with sensitive data attacker can trigger false positive • [secret] + [user input] => auditor fired? attacker may leak [secret] by changing [user input] • Antivirus can be an oracle as well! Various analyzers for contents

-

Abusing Antivirus Technologies • Antivirus Software is blackbox When

they work? What they will do? How they detect malware? How is their architecture? Which files are required to run them? etc. • Let's dig into Windows Defender Most popular Running on Windows by default

-

Windows Defender • What content will be detected as malicious?

They have their own malware list https://www.microsoft.com/en-us/wdsi/definitions/antimalware-definition-release-notes Probably other vendors have similar ones. No details published • We need to analyze Windows Defender!

-

Black-box Windows Defender analysis • Run audit process on...

file access command execution if (malicious) block access from user and notify to user • Analyzers for various content Encoding Base64 Archive, Compression ZIP, GZip, ... Executables PE, WSH (VBS, JScript), ... • Black-box analyzing is super tiresome work

-

How to analyze Windows Defender efficiently?

-

Windows Defender analysis is tiresome work • "MpCmdRun.exe" can trigger

the engine directly still some issues are there

-

Windows Defender analysis is tiresome work • Unexpected behavior of

Windows Defender timing issue neutralization (deletion) • We have to regenerate payloads ...bunch of times • No debug information Hard to know why one is detected or not detected • Any useful tools? several works are there

-

Windows Defender is ported to Linux! • github.com/taviso/loadlibrary emulating

mpengine.dll execution enables us to do try and error show us some debug output ~$ ./mpclient ../files/eicar main(): Scanning ../files/eicar... EngineScanCallback(): Scanning input EngineScanCallback(): Threat Virus:DOS/EICAR_Test_File identified. ~$ ./mpclient ../files/eicar.b64 main(): Scanning ../files/eicar.b64... EngineScanCallback(): Scanning input EngineScanCallback(): Scanning input->(Base64) EngineScanCallback(): Threat Virus:DOS/EICAR_Test_File identified.

-

Some tips about taviso/loadlibrary • You can get PDB symbol

file in older version refer github.com/0xAlexei/WindowsDefenderTools MD5=e95d3f9e90ba3ccd1a4b8d63cbd88d1b => 1.271.81.0 Download older version of mpam-fe.exe then use cabextract mpengine.dll is core engine • Debug features enable DEBUG flag to trace API calls inside

-

Windows Defender internals • Windows Defender signature format: *.vdm

mpasbase.vdm somehow encrypted • WDExtract would be helpful github.com/hfiref0x/WDExtract • Let's see the contents decrypted

-

Windows Defender internals • Windows Defender uses Lua

-

Windows Defender internals signature name signature definition (string)

-

Windows Defender internals • handlers for various file format

-

After white-box (?) Windows Defender analysis • Windows Defender has

JScript analyzer with DOM API supported • Not just parsing, but also emulating • If JScript calls eval(str) , str would also be audited eval("EICAR") => detected • What happens if combined

-

Attack to demo application • Simple application for PoC

GET /?c1=controllable1&c2=controllable2 save data with simple format user cannot see the content of Secret GET /:name check existence and integrity • How to leak the Secret ?

-

Building exploit • We have Windows Defender emulator! with

debug information • JScript eval function also evaluates argument threat detected if argument contains malicious • eval("EICA" + input) => ? threat detected input is "R" nothing detected input is not "R"

-

Some issues in JScript engine • if statement will never

be evaluated if (true) {eval("EICA" + "R")} not detected object accessing will help you: {0: "a", 1: "b", ...}[input] • parser stops on null byte eval("EICA" + "R[NULL]") syntax error how to deal with null bytes?

-

Another feature in mpengine.dll • They can analyze HTML document

some html tags would be a trigger (ex. <script>) parser will not stop on null byte • JScript can access the elements :) if they have <body> tag <script>document.body.innerHTML[0]</script> <body>[secret]</body> • Now you have an oracle!

-

Building exploit • JavaScript $idx and $c would be

iterated • Windows Defender get angry if $c is appropriate • It requires 256 times try for each $idx : (var body = document.body.innerHTML; var eicar = "EICA"; var n = body[$idx].charCodeAt(0); eicar = eicar + String.fromCharCode(n^$c); eval(eicar);

-

Building exploit • much more faster! Math.min is also

available, do binary search • $c < [input]: detected • $c > [input]: not detected then do binary search! var body = document.body.innerHTML; var eicar = "EICA"; var n = body[$idx].charCodeAt(0); eicar = eicar + {$c: 'k'}[Math.min($c, n)]; eval(eicar);

-

Building exploit • Now everything is ready :) Controllable1:

<script>...</script><body> Secret: [secret] Controllable2: </body> • to get oracle: accessing /:name after querying / detected Internal Server Error not detected you can see the response ...<script>[script]</script><body>...[secret]...</body>...

-

Demo • AVOracle attack against simple demo application

-

Pros and Cons • Pros Attacker can use this

method blindly No need to know target structure well, just put part of payloads everywhere • Cons Attacker need to put two pieces of payloads Only data between payloads would be leaked • Any other variants? It would be great if there is way to leak previous / following data No PoC so far

-

Any other victims?

-

Potential victims • So many applications are saving user input

with sensitive data • Session file TokyoWesterns CTF 2019 phpnote leak HMAC secret stored in PHP session (not visible from user) • Log file Apache, Nginx, IIS • Database file-based DBMS (ex. SQLite3) • Cache file browser, byte code cache

-

Antivirus as file modifier • Antivirus deletes / modifies file

if detected Windows Defender replaces matched part by spaces (in case of HTML script tag) • Attacker can delete content partially • Even attacker cannot leak data, there might be something to do data<script>eval('EICAR');</script>data data<script> </script>data

-

Wiping out evidence • Attacker can delete part of log

1. put <script>/* before beginning attack 2. do malicious attempts 3. put */;eval('EICAR');</script> *after attack x.x.x.x - - [29/Oct/2019:00:00:00 +0000] "GET /<script>/*</script>*" x.x.x.x - - [29/Oct/2019:00:00:10 +0900] "GET /attack.php" ... [some malicious attempts] ... x.x.x.x - - [29/Oct/2019:00:00:00 +0000] "GET /*/;eval('EICAR')</script>"

-

Wiping out evidence • Attacker can delete part of log

1. put <script>/* before beginning attack 2. do malicious attempts 3. put */;eval('EICAR');</script> after attack x.x.x.x - - [29/Oct/2019:00:00:00 +0000] "GET /<script> </script>"

-

Antivirus as DoS • Deleting matched malicious over structural boundary

• Structure metadata would be destroyed replaced by spaces • If two part will not over the boundary attacker can overwrite other data

-

How about other antivirus?

-

Targeting other antivirus • VirusTotal is the best friend :)

• Which antivirus suppoorts JScript emulator? eval('EICA'+'R'); // should be detected eval('EICA'+'#'); // should not be detected • 4 antivirus passed Cyren DrWeb Microsoft NANO-Antivirus • TrendMicro false positive

-

Targeting other antivirus • Further testing • Which antivirus supports

DOM API? eval('EICA'+innerHTML[0]);<body>R</body> // should be detected eval('EICA'+innerHTML[0]);<body>#</body> // should not be detected • Only Microsoft passed That's why they are vulnerable to AVOracle • SUPERAntiSpyware false positive

-

Windows Defender is too smart

-

How to prevent this attack? • IMO: no generic way

to patch standard behavior, not vulnerability • Disable auditor engine is one way Chromium XSS auditor is removed but Microsoft would not remove the engine • Application developers should ... know about this attack not save secret with controllable data • ... but it is not developer's fault! Antivirus vendor should take care about that

-

Conclusions • Auditor Based Oracle is everywhere Antivirus is

one big example it would be oracle if it has intelligent engine • Windows Defender is too smarter than other antivirus resulted in an effective oracle more smarter engine will get more oracles • Antivirus behavior would be sometimes harmful not only data leakage, also DoS • DO NOT store any secret surrounded by user input or your application would be vulnerable to AVOracle

-

Any Questions? @t0nk42 icchy