The development of a contained and user emulated malware assessment platform

Siebe Hodzelmans & Frank Potter
Google Acquires Online Virus, Malware and URL Scanner VirusTotal

Frederic Lardinois  @fredericl  /  6 years ago

VirusTotal, an online malware and virus scanner, was just acquired by Google. The company already used a number of Google services ahead of the acquisition, including App Engine and Google Storage. VirusTotal will continue to operate

(TechCrunch, 2012)
Incident Response and Malware Analysis

(Debrie, Lone-Sang, and Quint, 2014)
Kaspersky: Yes, we obtained NSA secrets. No, we didn’t help steal them

Moscow-based AV provider challenges claims it helped Russian spies.

DAN GOODIN - 11/16/2017, 11:00 AM

(ArsTechnica, 2017)
What Do Antivirus’s Actually Do?

Antivirus software is one of those things we can often take for granted. Many people have a “set it and forget it” approach to their antivirus settings. Or they may be under the misconception that antivirus software simply scans your files for viruses.

(Times Square Chronicles, 2019)
Research question

‘How can malware be tested for detection of antivirus software by emulating user actions, without the AV vendor learning about the malware?’
Sub questions

● What traffic is generated by AV software?

● How to prevent AV software from notifying and submitting the red team’s malware to the AV vendor?

● Are there any differences between direct scanning and user emulated detection rates?
Methodology - Traffic analysis

- McAfee, Symantec and Trend Micro
- Malware samples
1. Client sends HTTPS CONNECT
2. Proxy sends HTTPS ESTABLISHED
3. Client initiates SSL handshake
4. Proxy initiates SSL handshake
5. Server responds with Certificate S
6. Proxy responds with Certificate P and completes SSL handshake
7. Client sends HTTPS REQUEST
8. Proxy sends HTTPS REQUEST
9. Server sends HTTPS RESPONSE
10. Proxy sends HTTPS RESPONSE
Methodology - Preventing submission
Whitelisting

Windows Client

Allow traffic flow

Allow traffic flow

Internet

Blacklisting

Windows Client

Block traffic flow

Block traffic flow

Block traffic flow

Block traffic flow

Internet

Offline

Windows Client

Snapshot

VM

Restore

VM

Internet
Methodology - User emulation

- Compare manual with emulated behavior of malware
- Web browsing user emulation with pywinauto and pyautogui
- Malware infection Tree
Results - Traffic analysis

● Traffic capture:
  ○ McAfee, Symantec and Trend Micro
  ○ Later Kaspersky

● In general:
  ○ Installation, registration, updating
  ○ Analytical data
  ○ Lots of hashes and encoded data
  ○ Only HTTP(S)
Results - Traffic analysis

● Noteworthy:
  ○ Trend Micro: missing SNI, long plain HTTP GET
  ○ McAfee: every file gets hashed, google analytics
  ○ Symantec: ping submission with data buffer
  ○ Kaspersky: lot of HTTP(S) 400 and 502 errors, certificate pinning

● No sample submission
## Results - Traffic analysis

<table>
<thead>
<tr>
<th></th>
<th>Das Malwerk</th>
<th>Deloitte obfuscated</th>
<th>Deloitte direct exports</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1e84-ff45</td>
<td>1f7b-55c7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>230a-6f87</td>
<td>266a-11f5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2578-6c51</td>
<td>obf. exe</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>obf. dll 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>obf. dll 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>beacon exe</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>beacon dll</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>msf vnm</td>
<td></td>
</tr>
<tr>
<td>McAfee</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Symantec</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Trend Micro</td>
<td>✗, ✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Kaspersky</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
</tbody>
</table>
Results - Sample submission prevention

- Offline: undesirable
  - Difference Trend Micro
  - Warning Symantec
- Blacklisting
  - Unsure what to block
  - Updates can change endpoints
- Whitelisting
  - Robust
  - Parameters:
    - hostnames
    - traffic size and direction
    - content

![Limited Scan Available](image_url)
Results - User emulation

- Two ways:
  - pywinauto, accessibility API
  - pyautogui, mouse and keyboard, screenshots
- Compared manual to emulation
  - Malware infection Tree
  - File handles, process tree structure
Results - User emulation
Results - User emulation

**Static malware analysis**

- Symantec: 100 (True positive), 47.4 (False positive)
- TrendMicro: 89 (True positive), 37 (False positive)
- McAfee: 52.6 (True positive), 26.3 (False positive)

**Dynamic malware analysis**

- Symantec: 78.9 (True positive), 0 (False positive)
- TrendMicro: 78.9 (True positive), 0 (False positive)
- McAfee: 15.8 (True positive), 0 (False positive)
Discussion

- Contamination of packet captures
- mitmproxy
  - Insecure connections
  - Kaspersky errors
- Results of sample submission prevention
  - Unable to trigger sample submission
  - Flaw in research design
  - Based on what we did observe
- McAfee low detection rate
Conclusion

- Variety of traffic
  - But no sample submission
- Whitelisting the best approach
- Dynamic analysis is of added value
  - User emulation matches manual
  - Multiple approaches to emulation

How can malware be tested for detection of antivirus software by emulating user actions, without the AV vendor learning about the malware?
Future work

- Exploratory investigation in traffic generated by AV software
  - Another approach: reverse engineering
- Combine whitelisting with IRMA
- Monitoring AV detection of malware
Questions?