When Malware Changed Its Mind: An Empirical Study of Variable Program Behaviors in the Real World

Erin Avllazagaj¹, Ziyun Zhu², Leyla Bilge³, Davide Balzarotti⁴, Tudor Dumitraş¹

¹University of Maryland, College Park
²Facebook
³NortonLifeLock Research Group
⁴EURECOM
Malware Behavior Changes Across Environments
Malware Behavior Changes Across Environments

• Missing libraries, different language settings, etc.\(^1\)

Malware Behavior Changes Across Environments

• Missing libraries, different language settings, etc.\(^1\)

• Prudent practices\(^2\):
  – “[…] caution generalizing from a single OS version […]”

---

Malware Behavior Changes Across Environments

• Example: Ramnit Worm

```c
int __cdecl try_to_exploit(LPSTR lpCommandLine)
{
    if ( !is_win() && !is_win8_1() )
    {
        if ( is_xp() )
        {
            if ( !check_updates_xp((int)"KB3000061") )
            {
                if ( is_admin() )
                    return 1;
            }
            LABEL_6:
            executeCVE_2014_4113(lpCommandLine);
            return 1;
        }
        }
    else if ( !check_updates_other((int)"KB3000061") )
    {
        if ( is_admin() && check_authority() > 1 )
            return 1;
        goto LABEL_6;
    }
    try_second_exploit(lpCommandLine);
    return 1;
}
return 0;
```

Malware Behavior Changes Across Environments

• Example: **Ramnit Worm**
  – Exploits CVE-2013-3660
    • Line 22
    • Local Privilege escalation on Win 7
  • Creates hundreds of **mutexes**
    – until exploit succeeds

```
int __cdecl try_to_exploit(LPSTR lpCommandLine)
{
    if (!is_win8() && !is_win8_1())
    {
        if (is_xp())
        {
            if (!check_updates_xp((int)"KB3000061") )
            {
                if (is_admin())
                    return 1;
            }
        }
    }
    execute_CVE_2014_4113(lpCommandLine);
    return 1;
}
else if (!check_updates_other((int)"KB3000061") )
{
    if (is_admin() && checkAuthority() > 1 )
        return 1;
    goto LABEL_6;
}
try_second_exploit(lp CommandLine);
return 1;
return 0;
```

Malware Behavior Changes Across Environments

• Example: **Ramnit Worm**
  - Exploits CVE-2013-3660
    • Line 22
    • Local Privilege escalation on Win 7
    • Creates hundreds of **mutexes**
      - until exploit succeeds
  - Only works on:
    • vulnerable OS versions
    • when run in **non-admin**

```c
int __cdecl try_to_exploit(LPSTR lpCommandLine)
{
    if ( !is_win8() && !is_win8_1() )
    {
        if ( is_xp() )
        {
            if ( !check_updates_xp((int)"KB3000061") )
            {
                if ( is_admin() )
                    return 1;

            LABEL_6:
            execute_CVE_2014_4113(lpCommandLine);
            return 1;
        }    
    }
    
    else if ( !check_updates_other((int)"KB3000061") )
    {
        if ( is_admin() && check_authority() > 1 )
            return 1;
        goto LABEL_6;
    }    
    
    try_second_exploit(lpCommandLine);
    return 1;
}
return 0;
```

Research Questions

RQ1: Variability analysis in the wild

– What parts of the execution trace vary more? And by how much?
Research Questions

**RQ1**: Variability analysis in the wild
- What parts of the execution trace vary more? And by how much?

**RQ2**: Invariant analysis in the wild
- Can we find behavioral invariants across executions?
Research Questions

**RQ1: Variability analysis in the wild**
- What parts of the execution trace vary more? And by how much?

**RQ2: Invariant analysis in the wild**
- Can we find behavioral invariants across executions?

**RQ3: Impact of variability**
- What is the impact of variability on malware detection and clustering?
The Dataset

- **7.6M** execution traces
- **5.4M** real users' machines in **>100 countries** in the world
- From **2018**
The Dataset

• **7.6M** execution traces
• **5.4M** real users' machines in **>100 countries** in the world
• From **2018**
• No private data is collected, passive recording
The Dataset (introduction)

<table>
<thead>
<tr>
<th>Action type</th>
<th>File name</th>
<th>File path</th>
<th>...</th>
<th>sample Hash</th>
<th>Exec.id</th>
<th>Mach.id</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Create</td>
<td>setup.exe</td>
<td>CSDL_PROFILE</td>
<td>...</td>
<td>AAA</td>
<td>1</td>
<td>abc</td>
</tr>
<tr>
<td>Mtx. Create</td>
<td>mtx!asjkf</td>
<td></td>
<td>...</td>
<td>ABC</td>
<td>5243523</td>
<td>abd</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
## The Dataset (introduction)

<table>
<thead>
<tr>
<th>Action type</th>
<th>File name</th>
<th>File path</th>
<th>...</th>
<th>sample Hash</th>
<th>Exec.id</th>
<th>Mach.id</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Create</td>
<td>setup.exe</td>
<td>CIDL_PROFILE</td>
<td>...</td>
<td>AAA</td>
<td>1</td>
<td>abc</td>
</tr>
<tr>
<td>Mtx. Create</td>
<td>mtx!asjkf</td>
<td>-</td>
<td>...</td>
<td>ABC</td>
<td>5243523</td>
<td>abd</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

When Malware Changed Its Mind

### (split by hash)

<table>
<thead>
<tr>
<th>Action type</th>
<th>File name</th>
<th>File path</th>
<th>...</th>
<th>sample Hash</th>
<th>Exec.id</th>
<th>Mach.id</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Create</td>
<td>setup.exe</td>
<td>CIDL_PROFILE</td>
<td>...</td>
<td>AAA</td>
<td>1</td>
<td>abc</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>AAA</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Action type</th>
<th>File name</th>
<th>File path</th>
<th>...</th>
<th>sample Hash</th>
<th>Exec.id</th>
<th>Mach.id</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mtx. Create</td>
<td>mtx!asjkf</td>
<td>-</td>
<td>...</td>
<td>ABC</td>
<td>5243 523</td>
<td>abd</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>ABC</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
## The Dataset (introduction)

### Action type | File name | File path | sample Hash | Exec.id | Mach.id
---|---|---|---|---|---
File Create | setup.exe | CSIDL_PROFILE | AAA | 1 | abc
Mt. Create | mtx!asjkf | - | ABC | 5243523 | abd

... | ... | ... | ... | ... | ...

Using VirusTotal labels and AVClass\(^1\) (2019) we found:

22K benign, 2.4K malware and 1.6K PUP

---

### The Dataset (introduction)

<table>
<thead>
<tr>
<th>Action type</th>
<th>File name</th>
<th>File path</th>
<th>... sample Hash</th>
<th>Exec.id</th>
<th>Mach.id</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Create</td>
<td>setup.exe</td>
<td>CSIDL_PROFILE</td>
<td>... AAA</td>
<td>1</td>
<td>abc</td>
</tr>
<tr>
<td>Mtx. Create</td>
<td>mtx!asjkf</td>
<td>-</td>
<td>... ABC</td>
<td>5243523</td>
<td>abd</td>
</tr>
</tbody>
</table>

When Malware Changed Its Mind

Using VirusTotal labels and AVClass\[1\] (*2019*)

we found:

#### 22K benign, 2.4K malware and 1.6K PUP

---

RQ1: Variability Analysis In The Wild

- Ramnit worm exploit

```c
int __cdecl try_to_exploit(LPSTR lpCommandLine)
{
    if ( !_is_win8() && !_is_win8_1() )
    {
        if ( _is_xp() )
        {
            if ( !_check_updates_xp(int)"KB3000061") )
            {
                if ( _is_admin() )
                    return 1;
            }
        }
    }
    else if ( !_check_updates_other(int)"KB3000061") )
    {
        if ( _is_admin() && _check_authority() > 1 )
            return 1;
        goto LABEL_6;
    }
    try_second_exploit(lpCommandLine); return 1;
}
```

RQ1: Variability Analysis In The Wild

• Ramnit worm exploit

• When this line is reached
  – ~100 more mutex create events
  – based on the **machine**

```c
int __cdecl try_to_exploit(LPSTR lpCommandLine)
{
    if (!is_win() && !is_win_xp())
    {
        if (is_xp())
        {
            if (!check_updates_xp((int)"KB300061") )
            {
                if (is_admin() )
                    return 1;
            }
        }
        goto LABEL_6;
        execute_CVE_2014_4113(lpCommandLine);
        return 1;
    }
    else if (!check_updates_other((int)"KB3000061") )
    {
        if (is_admin() && check_authority() > 1 )
            return 1;
        goto LABEL_6;
    }
    try_second_exploit(lpCommandLine);
    return 1;
}
```

### Methodology
(for each hash)

<table>
<thead>
<tr>
<th>Action type</th>
<th>File name</th>
<th>File path</th>
<th>samp. Hash</th>
<th>Exec.id</th>
<th>Mach. id</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Create</td>
<td>setup.exe</td>
<td>CSIDL_PR FILE</td>
<td>AAA</td>
<td>1</td>
<td>abc</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>... AAA</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
### Methodology (for each hash)

When Malware Changed Its Mind 20

<table>
<thead>
<tr>
<th>Action type</th>
<th>File name</th>
<th>File path</th>
<th>…</th>
<th>samp le Hash</th>
<th>Exec.i d</th>
<th>Mach. id</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Create</td>
<td>setup.exe</td>
<td>CSDL_PR OFILE</td>
<td>…</td>
<td>AAA</td>
<td>1</td>
<td>abc</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>AAA</td>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>

(Group by machine ID and remove executions after week 0)

<table>
<thead>
<tr>
<th>Action type</th>
<th>File name</th>
<th>File path</th>
<th>…</th>
<th>samp le Hash</th>
<th>Exec.i d</th>
<th>Mach. id</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Create</td>
<td>2222.exe</td>
<td>CSDL_PR OFILE</td>
<td>…</td>
<td>AAA</td>
<td>4</td>
<td>aaa</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>AAA</td>
<td>3</td>
<td>aaa</td>
</tr>
</tbody>
</table>

RQ1: Variability Analysis In The Wild (machines)
### RQ1: Variability Analysis In The Wild (machines)

#### Methodology (for each hash)

<table>
<thead>
<tr>
<th>Action type</th>
<th>File name</th>
<th>File path</th>
<th>…</th>
<th>samp le Hash</th>
<th>Exec.i d</th>
<th>Mach. id</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Create</td>
<td>setup.exe</td>
<td>CSIDL_PR OFILE</td>
<td>…</td>
<td>AAA</td>
<td>1</td>
<td>abc</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>AAA</td>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>

- **File Creations:** 5
- **Mutex Creations:** 2
- **Total:** 52

(Group by machine ID and remove executions after week 0)

<table>
<thead>
<tr>
<th>Action type</th>
<th>File name</th>
<th>File path</th>
<th>…</th>
<th>samp le Hash</th>
<th>Exec.i d</th>
<th>Mach. id</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Create</td>
<td>2222.exe</td>
<td>CSIDL_PR OFILE</td>
<td>…</td>
<td>AAA</td>
<td>4</td>
<td>aaa</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>AAA</td>
<td>3</td>
<td>aaa</td>
</tr>
</tbody>
</table>

- **File Creations:** 5
- **Mutex Creations:** 42
- **Total:** 92
RQ1: Variability Analysis In The Wild (machines)

Methodology (for each hash)

<table>
<thead>
<tr>
<th>Action type</th>
<th>File name</th>
<th>File path</th>
<th>...</th>
<th>samp le Hash</th>
<th>Exec.i d</th>
<th>Mach. id</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Create</td>
<td>setup.exe</td>
<td>CSIDL_PR OFILE</td>
<td>...</td>
<td>AAA</td>
<td>1</td>
<td>abc</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>AAA</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

(Group by machine ID and remove executions after week 0)

File Creations: 5
Mutex Creations: 2
Total: 52

[ 45, ... , 52 , ... , 92 , ... , 100 ]

IQR → 92 − 52 = 40

IQR → 92 − 52 = 40
RQ1: Variability Analysis In The Wild (machines)

Methodology (for each hash)

<table>
<thead>
<tr>
<th>Action Type</th>
<th>File Name</th>
<th>File Path</th>
<th>Sample Hash</th>
<th>Exec.i.d</th>
<th>Mach. id</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Create</td>
<td>setup.exe</td>
<td>CSIDL_PROFILE</td>
<td>AAA</td>
<td>1</td>
<td>abc</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Group by machine ID and remove executions after week 0)

File Creations: 5
Mutex Creations: 2

Total: 52

Analysis in the paper

<table>
<thead>
<tr>
<th>Action Type</th>
<th>File Name</th>
<th>File Path</th>
<th>Sample Hash</th>
<th>Exec.i.d</th>
<th>Mach. id</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Create</td>
<td>2222.exe</td>
<td>CSIDL_PROFILE</td>
<td>AAA</td>
<td>4</td>
<td>aaa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

File Creations: 5
Mutex Creations: 42

Total: 92

[ 45, ..., 52, ..., 92, ..., 100 ]

IQR \( \rightarrow 92 - 52 = 40 \)
RQ1: Variability Analysis In The Wild (machines)

\[ \text{IQR} \rightarrow 92 - 52 = 40 \]
\[ \text{IQR} \rightarrow 10 \]
\[ \text{IQR} \rightarrow 0 \]
\[ \text{IQR} \rightarrow 100 \]
\[ \text{IQR} \rightarrow 60 \]

For all malware (blue boxplot)

For all PUP (orange boxplot)

For all benigns (green boxplot)
RQ1: Variability Analysis In The Wild (machines)

- At least 50% of the malware:
  - 59 missing or additional actions
RQ1: Variability Analysis In The Wild (machines)

- At least 50% of the malware:
  - **59** missing or additional actions

- File creation
  - The major source of machine-induced variability in malware.
RQ1: Variability Analysis In The Wild (machines)

- Methodology:
  - IQR of the number of unique parameter values across different machines.

- Number of unique file names varies by 25 across machines

<table>
<thead>
<tr>
<th>File</th>
<th>Median</th>
<th>75th percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mal</td>
<td>PUP</td>
</tr>
<tr>
<td>Path</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Name</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>Ext.</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>
RQ1 Summary

• High variability in malware across machines
  – File Creation makes up most of variability in malware
  – File name is the most variable parameter
RQ2: Invariant Analysis In The Wild

- Focus on action-parameter pair
  - used in Sigma
  - used in cuckoo

```python
class CreatesUserFolderEXE(Signature):
    name = "creates_user_folder.exe"
    description = "Creates an executable file in a user folder"
    severity = 3
    families = ["persistance"]
    authors = ["Kevin Ross"]
    minimum = "2.0"
    tto = ["T1129"]

directories_re = {
    "([a-zA-Z]):\Users\([^\\]+\)\AppData\[^\\]+\.*",
    "([a-zA-Z]):\Documents\ and\ Settings\[^\\]+\[^\\]+\Local\ Settings\[^\\]+\.*",
}

def on_complete(self):
    for dropped in self.get_results("dropped", []):
        if "filepath" in dropped:
            droppedtype = dropped["type"]
            filepath = dropped["filepath"]
            if "MS-DOS executable" in droppedtype:
                for directory in self.directories_re:
                    if re.match(directory, filepath):
                        self.mark_loc("file", filepath)

return self.has_marks()
```

https://github.com/cuckoosandbox/community/tree/master/modules/signatures

logsource:
category: process_creation
product: windows
detection:
selection:
  CommandLine: \\
  condition: selection
  https://github.com/SigmaHQ/sigma

https://github.com/SigmaHQ/sigma
RQ2: Invariant Analysis In The Wild

• Focus on action-parameter pair
  – used in Sigma
  – used in cuckoo

```python
class CreatesUserFolderEXE(Signature):
    name = "creates_user_folder_exe"
    description = "Creates an executable file in a user folder"
    severity = 3
    families = ["persistance"]
    authors = ["Kevin Ross"]
    minimum = "2.0"
    tto = ["T1129"]

directories_re = [
    "\[a-zA-Z]:\\\([\:\:\:\:\:\:\:\:\:\\]\\\:\\:\:\:\:\:\:\\:*",
    "\[a-zA-Z]:\\\Documents\\\:and\\\:Settings\\\([\:\:\:\:\:\:\:\:\:\\]\\\:Local\\\:Settings\\\:*",
]

def on_complete(self):
    for dropped in self.get_results("dropped", []):
        if "filepath" in dropped:
            droppedtype = dropped["type"]
            filepath = dropped["filepath"]
            if "MS-DOS executable" in droppedtype:
                for directory in self.directories_re:
                    if re.match(directory, filepath):
                        self.mark_loc("file", filepath)

return self.has_marks()
```

https://github.com/cuckoosandbox/community/tree/master/modules/signatures

logsource:
  category: process_creation
  product: windows

detection:
  selection:
    CommandLine: "\*non\-\*bypass $*"
  condition: selection

https://github.com/SigmaHQ/sigma

When Malware Changed Its Mind
RQ2: Invariant Analysis In The Wild

Extract parameter values

Split them by delimiter

CSIDL_PROFILE
cmd.exe
del
virus.exe

mtx!asjkf
CSIDL_PROFILE/folder1
runprogram.exe
icon.png
CSIDL_APPDATA/bin
config.ini
setup.exe
RQ2: Invariant Analysis In The Wild

(Remove values that appear in benign samples)

wnry → appears in 30/50 machines → 60% → appear in 65% of the machines
virus → appears in 10/50 machines → 20%
RQ2: Invariant Analysis In The Wild

• How to maximize coverage?
  – **Assumption:** Sandbox is undetectable.
RQ2: Invariant Analysis In The Wild

• How to maximize coverage?
  – **Assumption**: Sandbox is undetectable.

• Pick \( n \) machines to get the bag of tokens
  – Check how much coverage would we get on the other machines.
RQ2: Invariant Analysis In The Wild

• How to maximize coverage?
  – Maximum coverage in 3 randomly generated machines
    • For file name tokens
  – One file name token doesn’t appear in all machines.
    • Use more than 1 file name

When Malware Changed Its Mind 35

~87%
RQ3: Impact Of Variability

• In terms of:
  – Clustering
  – Anomaly detection (AccessMiner\textsuperscript{[1]})

RQ3: Impact Of Variability (clustering)

• Methodology:
  – Get 4 executions per malware sample in the same week
  – Reproduce the clustering by Bailey et al.\textsuperscript{[1]}

[1] Bailey et al., Automated Classification and Analysis of Internet Malware, RAID 2007
RQ3: Impact Of Variability (clustering)

• Results (out of 2424 malware samples):
  – 1,624 (67%) in the same cluster
  – 655 (27%) in 2 clusters
  – 121 (5%) in 3 clusters
  – 24 (1%) in 4 different cluster
RQ3: Impact of variability (clustering)

• Results (out of 2424 malware samples):
  – 1,624 (67%) in the same cluster
  – 655 (27%) in 2 clusters
  – 121 (5%) in 3 clusters
  – 24 (1%) in 4 different cluster

clustering results with 1 trace per sample may not correctly cluster malware into families
Conclusions
Conclusions

• First measurement of malware behavior at scale:
  – Single trace per malware sample is not enough
Conclusions

• First measurement of malware behavior at scale:
  – Single trace per malware sample is not enough

• Variability in malware is greater than PUP and benign
  – Across both time and machines
Conclusions

• First measurement of malware behavior at scale:
  – Single trace per malware sample is not enough

• Variability in malware is greater than PUP and benign
  – Across both time and machines

• It’s still feasible to find invariant in malware behavior
  – AV vendors can safely do it
Thank you!

Erin Avllazagaj
albocode@umd.edu