Browser forensics: Adblocker extensions

Willem Rens (UvA MSc SNE student)

Supervisor: Johannes de Vries (Fox-IT)
Why traditional browser forensics may not work

- Cleared
  - Cookies
  - Cache
  - History

Why traditional browser forensics may not work

- **Cleared**
  - Cookies
  - Cache
  - History


- **Private browsing**
  - Incognito (Chrome)
  - InPrivate (Ie&edge)
  - Private browsing (Firefox)
Why traditional browser forensics may not work

- Cleared
  - Cookies
  - Cache
  - History


- Private browsing
  - Incognito (Chrome)
  - InPrivate (Ie&edge)
  - Private browsing (Firefox)

Claims to maintain complete user privacy by not storing traces of web browsing sessions. Flowers et al. (2016) studied the validity of this claim. IE11 still left traces, Chrome and Firefox did not.
Adblocker extension usage estimates

Usage estimates vary widely
Adblocker extension usage estimates

Usage estimates vary widely

● 20% ? (Metadata analysis within a large European ISP, 2015, Metwalley, et al.)
Ad blocker extension usage estimates

Usage estimates vary widely

- 20% ? (Metadata analysis within a large European ISP, 2015, Metwalley, et al.)
- 62% ? (Undergraduate business students, 2011, Sandvig, et al.)
Adblocker extension usage estimates

Usage estimates vary widely

- 20% ? (Metadata analysis within a large European ISP, 2015, Metwalley, et al.)
- 62% ? (Undergraduate business students, 2011, Sandvig, et al.)

41% increase year by year (Adobe and Pagefair, 2015)
Research questions

- RQ1 - What artifacts are stored by the tested ad-blocking extensions during *normal* and *private* browsing?
Research questions

- RQ1 - What artifacts are stored the tested ad-blocking extensions during *normal* and *private* browsing?

- RQ2 - If artifacts are found, what is their usefulness in browser forensics?
Tested browsers & their most popular Adblocker extension.

<table>
<thead>
<tr>
<th>Browser</th>
<th>Adblocker extension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozilla Firefox 46.0</td>
<td>Adblock Plus 2.8.2</td>
</tr>
<tr>
<td>Google Chrome/55.0.2883.87</td>
<td>AdBlock 3.8.4</td>
</tr>
<tr>
<td>Internet Explorer 11</td>
<td>Adblock Plus 1.6</td>
</tr>
<tr>
<td>Microsoft Edge/14.14393</td>
<td>AdBlock 1.9.0.0</td>
</tr>
</tbody>
</table>

AdBlock & Adblock Plus are not related.

Source most popular adblocking extensions = amount of downloads and reviews as stated by respective webstore. Other adblocking extensions have significant smaller market shares < 10%. 
Approach

- Automated sample gathering.
  - Control Sample.
  - Adblock Sample.
  - Private browsing sample.
  - Browsing session entails the visitation of top 50 NL websites as per alexa.com.
  - Python + selenium + save timestamps on url request.
  - Chrome & Firefox have the concept of user profiles, create a new one and extract the user data directory.
  - Ie & Edge more difficult to automate due to limited control with selenium, such as for adding an extension and it does not have the concept of user profiles.
Approach

● Automated sample gathering.
  ○ Control Sample.
  ○ Adblock Sample.
  ○ Private browsing sample.
  ○ Browsing session entails the visitation of top 50 NL websites as per alexa.com.
  ○ Python + selenium + save timestamps on url request.
  ○ Chrome & Firefox have the concept of user profiles, create a new one and extract the user data directory.
  ○ Ie & Edge more difficult to automate due to limited control with selenium, such as for adding an extension and it does not have the concept of user profiles.

● OSForensics (trialware)
  ○ Also used by Flowers et al. (2016).
  ○ Snapshots of the file system, compare them pre and after sample gathering.
Approach

- Automated sample gathering.
  - Control Sample.
  - Adblock Sample.
  - Private browsing sample.
  - Browsing session entails the visitation of top 50 NL websites as per alexa.com.
  - Python + selenium + save timestamps on url request.
  - Chrome & Firefox have the concept of user profiles, create a new one and extract the user data directory.
  - Ie & Edge more difficult to automate due to limited control with selenium, such as for adding an extension and it does not have the concept of user profiles.

- OSForensics (trialware)
  - Also used by Flowers et al. (2016).
  - Snapshots of the file system, compare them pre and after sample gathering.

- W10 Home 64-bit.
Approach

- **Automated sample gathering.**
  - *Control* Sample.
  - *Adblock* Sample.
  - *Private* browsing sample.
  - Browsing session entails the visitation of top 50 NL websites as per alexa.com.
  - Python + selenium + save timestamps on url request.
  - Chrome & Firefox have the concept of user profiles, create a new one and extract the user data directory.
  - Ie & Edge more difficult to automate due to limited control with selenium, such as for adding an extension and it does not have the concept of user profiles.

- **OSForensics (trialware)**
  - Also used by Flowers et al. (2016).
  - Snapshots of the file system, compare them pre and after sample gathering.

- **W10 Home 64-bit.**

- Research indicates 80% of software is used in its default setting, Wills et al. (2016) confirms this for the use of Adblock Plus.
Approach

- Automated sample gathering.
  - Control Sample.
  - Adblock Sample.
  - Private browsing sample.
  - Browsing session entails the visitation of top 50 NL websites as per alexa.com.
  - Python + selenium + save timestamps on url request.
  - Chrome & Firefox have the concept of user profiles, create a new one and extract the user data directory.
  - Ie & Edge more difficult to automate due to limited control with selenium, such as for adding an extension and it does not have the concept of user profiles.

- OSForensics (trialware)
  - Also used by Flowers et al. (2016).
  - Snapshots of the file system, compare them pre and after sample gathering.

- W10 Home 64-bit.
- Research indicates 80% of software is used in its default setting, Wills et al. (2016) confirms this for the use of Adblock Plus.

But first explore the mechanisms used by ad blocking extensions and study its source code.
Adblocker mechanics

● Filter lists
  ○ By far most popular is EasyList
  ○ Whitelist filters overrule
Adblocker mechanics

- Filter lists
  - By far most popular is EasyList
  - Whitelist filters overrule

- Blocking requests
  - Extensions can register content policies, they get called whenever the browser needs to load something.
  - If there is a filter hit do not request the resource.
Adblocker mechanics

- Filter lists
  - By far most popular is EasyList
  - Whitelist filters overrule

- Blocking requests
  - Extensions can register **content policies**, they get called whenever the browser needs to load something.
  - If there is a filter hit do not request the resource.

- Hiding elements
  - Some elements can not be blocked otherwise page won’t load.
  - Update **user style sheet** (overrides other styling) with styling > **display: none !important**
addUserCSS(subject, selectors.map(
    selector => selector + "{display: none !important;}
    
);
addUserCSS(subject, selectors.map(
    selector => selector + "\{display: none !important;\}"
).join("\n"));

if (!isPrivate(subject))
    port.emit("addHits", filters);
Extensions storing capabilities

- **SessionStorage** - stores data for one session (data is lost when the browser tab is closed).
Extensions storing capabilities

- **SessionStorage** - stores data for one session (data is lost when the browser tab is closed).

- **LocalStorage** - stores data with no expiration date.
Extensions storing capabilities

- **SessionStorage** - stores data for one session (data is lost when the browser tab is closed).

- **LocalStorage** - stores data with no expiration date.

This concept is used in all the tested browsers.
Comparing samples

```
import os

def compare_dir_layout(dir1, dir2):
    print('files in "' + dir2 + '" but not in "' + dir1 + '"')
    for (dirpath, dirnames, filenames) in os.walk(dir1):
        for filename in filenames:
            relative_path = dirpath.replace(dir1, '')
            if os.path.exists(dir2 + relative_path + '\' + filename) == False:
                print(relative_path, filename)

compare_dir_layout('Control', 'AdBlock')
```
Comparing file change differences of samples.
**Results**  Google Chrome/55.0.2883.87 + AdBlock

Chrome local storage for extensions -> LevelDB  *(key-value store written by Google)*

<table>
<thead>
<tr>
<th>Key</th>
<th>Value (contents)</th>
</tr>
</thead>
<tbody>
<tr>
<td>blockage_stats</td>
<td>Epoch installation time</td>
</tr>
<tr>
<td>file:pattern.ini</td>
<td>Filter list + subscription</td>
</tr>
<tr>
<td>next_ping_time</td>
<td>Sends user data to <a href="https://ping.getadblock.com/stats/">https://ping.getadblock.com/stats/</a> on given epoch time</td>
</tr>
<tr>
<td>pref:blocked_total</td>
<td>Total amount of filter hits since installation</td>
</tr>
<tr>
<td>pref:currentVersion</td>
<td>Version number</td>
</tr>
<tr>
<td>pref:notificationdata</td>
<td>Stats about the subscriptions, including when to check for updates.</td>
</tr>
<tr>
<td>pref:settings</td>
<td>Some settings</td>
</tr>
<tr>
<td>pref:total_pings</td>
<td>Total amount of pings</td>
</tr>
<tr>
<td>userid</td>
<td>Unique user ID</td>
</tr>
</tbody>
</table>
### Results

Google Chrome/55.0.2883.87 + AdBlock

Chrome local storage for extensions \(\rightarrow\) LevelDB (key-value store written by Google)

<table>
<thead>
<tr>
<th>Key</th>
<th>Value (contents)</th>
</tr>
</thead>
<tbody>
<tr>
<td>blockage_stats</td>
<td>Epoch installation time</td>
</tr>
<tr>
<td>file:pattern.ini</td>
<td>Filter list + subscription</td>
</tr>
<tr>
<td>next_ping_time</td>
<td>Sends user data to <a href="https://ping.getadblock.com/stats/">https://ping.getadblock.com/stats/</a> on given epoch time</td>
</tr>
<tr>
<td>pref:blocked_total</td>
<td>Total amount of filter hits since installation</td>
</tr>
<tr>
<td>pref:currentVersion</td>
<td>Version number</td>
</tr>
<tr>
<td>pref:notificationdata</td>
<td>Stats about the subscriptions, including when to check for updates.</td>
</tr>
<tr>
<td>pref:settings</td>
<td>Some settings</td>
</tr>
<tr>
<td>pref:total_pings</td>
<td>Total amount of pings</td>
</tr>
<tr>
<td>userid</td>
<td>Unique user ID</td>
</tr>
</tbody>
</table>
Results  
Microsoft Edge/14.14393 + AdBlock 1.9.0.0

Edge local storage for extensions -> .dat

<table>
<thead>
<tr>
<th>Key</th>
<th>Value (contents)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blockage_stats</td>
<td>Epoch time first filter hit+ total amount of filter hits since installation, split between 'total' and 'malware_total'.</td>
</tr>
<tr>
<td>Filter_lists</td>
<td>Pointing to filter lists location.</td>
</tr>
<tr>
<td>Last_subscriptions_check</td>
<td>Epoch time last time filters were updated</td>
</tr>
<tr>
<td>Next_ping_time</td>
<td>Sends user data to <a href="https://ping.getadblock.com/stats/">https://ping.getadblock.com/stats/</a> on given epoch time</td>
</tr>
<tr>
<td>Settings</td>
<td>Settings</td>
</tr>
<tr>
<td>Total_pings</td>
<td>Total amount of pings</td>
</tr>
<tr>
<td>Userid</td>
<td>Unique user ID</td>
</tr>
</tbody>
</table>
### Results

**Microsoft Edge/14.14393 + AdBlock 1.9.0.0**

**Edge local storage for extensions -> .dat**

<table>
<thead>
<tr>
<th>Key</th>
<th>Value (contents)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blockage_stats</td>
<td>Epoch time first filter hit+ total amount of filter hits since installation, split between 'total' and 'malware_total'.</td>
</tr>
<tr>
<td>Filter_lists</td>
<td>Pointing to filter lists location.</td>
</tr>
<tr>
<td>Last_subscriptions_check</td>
<td>Epoch time last time filters were updated</td>
</tr>
<tr>
<td>Next_ping_time</td>
<td>Sends user data to <a href="https://ping.getadblock.com/stats/">https://ping.getadblock.com/stats/</a> on given epoch time</td>
</tr>
<tr>
<td>Settings</td>
<td>Settings</td>
</tr>
<tr>
<td>Total_pings</td>
<td>Total amount of pings</td>
</tr>
<tr>
<td>Userid</td>
<td>Unique user ID</td>
</tr>
</tbody>
</table>
Results Internet Explorer 11 + Adblock Plus 1.6

Patterns.ini -> filter list subscription + filters

Settings.ini -> settings other than default

prefs.json -> notificationdata
Results

Internet Explorer 11 + Adblock Plus 1.6

Patterns.ini -> filter list subscription + filters

Settings.ini -> settings other than default

defs.json -> notificationdata
**Results**  Mozilla Firefox 46.0 + Adblock Plus 2.8.2

<table>
<thead>
<tr>
<th>Loc</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>/adblockplus/patterns.ini</td>
<td><strong>Filter hits including a hitCounter and lastHit parameter</strong> + filterList</td>
</tr>
<tr>
<td>/adblockplus/patterns-backup1.ini</td>
<td>If patterns.ini is full patterns-backup.ini is created with a number incrementing from 1.</td>
</tr>
<tr>
<td><code>AdBlock\extensions\{d10d0bf8-f5b5-c8b4-a8b2-2b9879e08c5d\}/</code></td>
<td>AdBlock application files</td>
</tr>
<tr>
<td><code>prefs.js</code></td>
<td>Ablock Plus settings that are different than default are added here</td>
</tr>
</tbody>
</table>

Location is relative to the data directory of the Firefox profile.
Results  Mozilla Firefox 46.0 + Adblock Plus 2.8.2

<table>
<thead>
<tr>
<th>Loc</th>
<th>/adblockplus/patterns.ini</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>contents including a hitCounter and lastHit parameter + filterList</td>
</tr>
<tr>
<td>/adblockplus/patterns-backup1.ini</td>
<td>If patterns.ini is full patterns-backup.ini is created with a number incrementing from 1.</td>
</tr>
<tr>
<td>AdBlock\extensions{d10d0bf8-f5b5-c8b4-a8b2-2b9879e08c5d}/</td>
<td>AdBlock application files</td>
</tr>
<tr>
<td>prefs.js</td>
<td>Ablock Plus settings that are different than default are added here</td>
</tr>
</tbody>
</table>

Location is relative to the data directory of the Firefox profile.
Patterns.ini

- Filter
- hitCount (amount of times this filter is activated)
- Last time this filter is activated in epoch time

[Filter]
text=@@|\|redditmedia.com/ads/display/$subdocument, domain=reddit.com
hitCount=2
lastHit=1484873352591

[Filter]
text=@@|\|engine.a.redditmedia.com/ados?$script, domain=redditmedia.com
hitCount=1
lastHit=1484873351895

[Filter]
text=@@|\|zkcdn.net/Advertisers/$image, domain=redditmedia.com
hitCount=1
lastHit=1484873352842

[Filter]
text=@@|\|zkcdn.net^$stylesheet, domain=redditmedia.com, script
hitCount=4
lastHit=1484873352678

[Filter]
text=@@|\|www.google.nl^$elemhide,~third-party
hitCount=2
lastHit=1484873283242

[Filter]
text=@@|\|www.google.ru^$elemhide,~third-party
hitCount=2
lastHit=1484873463316
PoC - test on top 500 sites per https://moz.com/top500

- Bigger sample (top 500 websites per https://moz.com/top500).
PoC - test on top 500 sites per https://moz.com/top500

- Bigger sample (top 500 websites per https://moz.com/top500).
- Use Firefox options to clear history / cookies / caches.
PoC - test on top 500 sites per https://moz.com/top500

- Bigger sample (top 500 websites per https://moz.com/top500).
- Use Firefox options to clear history / cookies / caches.

143 / 500 site visits left traces by filter hits.
## Conclusion

**RQ1**: What artifacts are stored by the tested ad-blocking extensions during normal and private browsing?

<table>
<thead>
<tr>
<th>Mode</th>
<th>AdBlock + Chrome&amp;Edge</th>
<th>Adblock Plus + Ie</th>
<th>Adblock Plus Firefox</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Settings, filterlist, total amount of filterHits</td>
<td>Settings, filterlists</td>
<td>Settings, filterlists, filter hits.</td>
</tr>
<tr>
<td>Private</td>
<td>Settings, filterlist, total amount of filterHits</td>
<td>Settings, filterlists</td>
<td>Settings, filterlists</td>
</tr>
</tbody>
</table>
Conclusion

RQ2: If artifacts are found, what is their usefulness in browser forensics?

- Total hitcount since installation -> useless.
Conclusion

**RQ2**: If artifacts are found, what is their usefulness in browser forensics?

- Total hitcount since installation -> useless.
- Filter hits -> useful.
- 143 / 500 traces in filter hits leading to last time visited.
- Firefox market share -> 10.4%.
- Estimated usage of Adblock Plus -> 20%.

$$0.104 \times 0.2 \times \frac{143}{500} = \text{Minimum of } \sim 0.6\%$$
Future work

- Improve PoC by parsing the filter hits in such a way that domains can be classified as in:
  - definitely visited
  - maybe visited
Future work

- Improve PoC by parsing the filter hits in such a way that domains can be classified as in:
  - definitely visited
  - maybe visited

- Other adblocking extensions have a much smaller market share. So might not be interesting to test them. Use Windows tool Process Explorer instead of OSforensics.