How Tracking Companies Circumvented Ad Blockers Using WebSockets

Muhammad Ahmad Bashir, Sajjad Arshad, Engin Kirda, William Robertson, Christo Wilson

Northeastern University
Online Tracking
Online Tracking

Surge in online advertising (internet economy)

- Ad networks pour in billions of dollars.
- Value for their investment?
  - Extensive tracking to serve targeted ads.
Online Tracking

Surge in online advertising (internet economy)

- Ad networks pour in billions of dollars.
- Value for their investment?
  - Extensive tracking to serve targeted ads.

User concern over tracking

- Led to the proliferation of ad blocking extensions
Online Tracking

Surge in online advertising (internet economy)

• Ad networks pour in billions of dollars.
• Value for their investment?
  • Extensive tracking to serve targeted ads.

User concern over tracking

• Led to the proliferation of ad blocking extensions

Ad networks fight back

• E.g Using anti ad blocking scripts
Google & Safari

- Google evaded Safari’s third-party cookie blocking policy (Jonathan Mayer)
- ... by submitting a form in an invisible iFrame
- Google was fined $22.5M by FTC
This Talk

How **Ad Networks** leveraged a bug in Chrome API to **bypass Ad Blockers** using **WebSockets**
This Talk

How Ad Networks leveraged a bug in Chrome API to bypass Ad Blockers using WebSockets

1. What caused this?

2. How this bug was leveraged by ad networks?
Web Sockets
Web Sockets

HTTP/S
Web Sockets

HTTP/S

request
response
Web Sockets

HTTP/S

Chatting App
Web Sockets

HTTP/S

Chatting App

request
response

anything new?
Web Sockets

HTTP/S

Chatting App

Web Socket

request
response

anything new?
Web Sockets

HTTP/S

Chatting App

Web Socket

- Both client and server can send/receive data
- This is a persistent connection
Web Sockets

HTTP/S

Chatting App

Web Socket

- Both client and server can send/receive data
- This is a persistent connection
Ad Blockers
Ad Blockers

- Chrome extension `chrome.webRequest` API
- Extension can inspect / modify / drop outgoing requests
Ad Blockers

- Chrome extension `chrome.webRequest` API
- Extension can inspect / modify / drop outgoing requests
Ad Blockers

- Chrome extension `chrome.webRequest` API
- Extension can inspect / modify / drop outgoing requests
Ad Blockers

- Chrome extension `chrome.webRequest` API
- Extension can inspect / modify / drop outgoing requests

http://cnn.com/logo.jpeg → webRequest API

Usually borrowed from EasyList

ABP Rule List
Ad Blockers

- Chrome extension `chrome.webRequest` API
- Extension can inspect / modify / drop outgoing requests

http://cnn.com/logo.jpeg

```
webRequest API

url

Usually borrowed from EasyList

ABP Rule List
```
Ad Blockers

- Chrome extension **chrome.webRequest** API
- Extension can inspect / modify / drop outgoing requests

![Diagram](http://cnn.com/logo.jpeg)
Ad Blockers

- Chrome extension `chrome.webRequest` API
  - Extension can inspect / modify / drop outgoing requests

```
http://cnn.com/logo.jpeg
```

```
webRequest API
```

```
url
```

```
Usually borrowed from EasyList
```

```
Rule List
```

```
ABP
```

6
Ad Blockers

- Chrome extension `chrome.webRequest` API
- Extension can inspect / modify / drop outgoing requests

![Diagram of webRequest API flow]

- `http://cnn.com/logo.jpeg`
- `webRequest API` module
- URL `url` is checked against `Rule List` usually borrowed from EasyList
- `webRequest API` sends the request to the website

6
Ad Blockers

- Chrome extension `chrome.webRequest` API
- Extension can inspect / modify / drop outgoing requests
Ad Blockers

• Chrome extension `chrome.webRequest` API
• Extension can inspect / modify / drop outgoing requests

![Diagram of ad blocking process]

- `http://cnn.com/logo.jpeg` usually borrowed from EasyList
- `http://doubleclick.com/s1.js`
Ad Blockers

- Chrome extension `chrome.webRequest` API
- Extension can inspect / modify / drop outgoing requests

![Diagram showing ad blockers in action](image-url)

**Rule List**
Usually borrowed from EasyList

**Examples**
- `http://cnn.com/logo.jpeg`
  - Accepted by the extension
- `http://doubleclick.com/s1.js`
  - Blocked by the extension

![Chrome extension icon](image-url)
Ad Blockers

• Chrome extension `chrome.webRequest` API
• Extension can inspect / modify / drop outgoing requests
AdBlock Evasion
AdBlock Evasion

- Bug in `webRequest API`
  - `ws/wss` requests did not trigger the API
AdBlock Evasion

- Bug in **webRequest API**
  - ws/wss requests did not trigger the API
AdBlock Evasion

- Bug in webRequest API
- ws/wss requests did not trigger the API
AdBlock Evasion

- Bug in \texttt{webRequest API}
- \texttt{ws/wss} requests did not trigger the API

Original bug reported

Users report unblocked ads

2012 2013 2014 2015 2016 2017 2018
AdBlock Evasion

• Bug in webRequest API

• ws/wss requests did not trigger the API

Original bug reported
2012 2013 2014

Users report unblocked ads
2015 2016

Patch Finalized (Landed)
2017 2018
AdBlock Evasion

- Bug in webRequest API
- ws/wss requests did not trigger the API

- Original bug reported:
  - 2012

- Users report unblocked ads:
  - 2015

- Patch Finalized (Landed):
  - 2016

- Chrome 58 released:
  - 2017

- 2018
AdBlock Evasion

- Bug in `webRequest API`
- `ws/wss` requests did not trigger the API

* Represents when our crawls were done

- Original bug reported
- `Users report unblocked ads`
- Patch Finalized (Landed)
- Chrome 58 released

2012 2013 2014 2015 2016 2017 2018
Data Crawling
Data Crawling

100K websites sampled from Alexa
Data Crawling

100K websites sampled from Alexa

Visit 15 links / website

Collected chains for all included resources
Data Crawling

- 100K websites sampled from Alexa
- Visit 15 links / website
- Collected chains for all included resources

This means we know which resource included which other resource
Data Crawling

- 100K websites sampled from Alexa
- Visit 15 links / website
- Collected chains for all included resources
  - Filter WebSockets
  - Filter all resources which end in web sockets
  - This means we know which resource included which other resource
Data Crawling

- 100K websites sampled from Alexa
  - Visit 15 links / website

- Collected chains for all included resources
  - Filter WebSockets
  - Filter all resources which end in web sockets
  - Detect A&A WebSockets
  - Mark web sockets which are used by A&A domains

This means we know which resource included which other resource

A&A = Advertising and Analytics
  - e.g. DoubleClick, Criteo, Adnxs
Data Crawling

100K websites sampled from Alexa

Visit 15 links / website

Collected chains for all included resources

Filter all resources which end in WebSockets

Filter WebSockets

Mark web sockets which are used by A&A domains

Detect A&A WebSockets

This means we know which resource included which other resource

Example Inclusion Tree

- pub/
  - index.html
    - srv.ws
    - ads/
      - script.js
      - ads/
        - frame.html
        - adnet/
          - data.ws
        - ads/
          - img_a.jpg

A&A = Advertising and Analytics
e.g. DoubleClick, Criteo, Adnxs
Data Crawling

100K websites sampled from Alexa

Visit 15 links / website

Collected chains for all included resources

Filter WebSockets

This means we know which resource included which other resource

Filter all resources which end in web sockets

Mark web sockets which are used by A&A domains

Detect A&A WebSockets

A&A = Advertising and Analytics e.g. DoubleClick, Criteo, Adnxs

Example Inclusion Tree

pub/index.html

- srv.ws
  - WebSocket

- ads/script.js
  - ads/frame.html

- adnet/data.ws
  - WebSocket

- ads/img_a.jpg

Example Inclusion Tree

WebSocket

WebSocket

WebSocket
Data Crawling

100K websites sampled from Alexa

Visit 15 links / website

Collected chains for all included resources

Filter all resources which end in web sockets

Filter WebSockets

Detect A&A WebSockets

This means we know which resource included which other resource

Mark web sockets which are used by A&A domains

A&A = Advertising and Analytics e.g. DoubleClick, Criteo, Adnxs

Example Inclusion Tree

pub/index.html

srv.ws

WebSocket

ads/script.js

ads/frame.html

adnet/data.ws

WebSocket

WebSocket
Data Crawling

100K websites sampled from Alexa

Visit 15 links / website

Collected chains for all included resources

Filter all resources which end in WebSockets

Filter WebSockets

Mark web sockets which are used by A&A domains

Detect A&A WebSockets

A&A = Advertising and Analytics
e.g. DoubleClick, Criteo, Adnxs

Example Inclusion Tree

pub/index.html

ads/script.js

ads/frame.html

adnet/data.ws

WebSocket

This means we know which resource included which other resource
High-Level Numbers
## High-Level Numbers

<table>
<thead>
<tr>
<th>Crawl Dates</th>
<th>% Websites with sockets</th>
<th>% Sockets with A&amp;A Initiators</th>
<th>% Sockets with A&amp;A Receivers</th>
<th>#Unique A&amp;A Initiators</th>
<th>#Unique A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 02-05, 2017</td>
<td>2.1</td>
<td>60.6</td>
<td>73.7</td>
<td>75</td>
<td>16</td>
</tr>
<tr>
<td>Apr 11-16, 2017</td>
<td>2.4</td>
<td>61.3</td>
<td>74.6</td>
<td>63</td>
<td>18</td>
</tr>
</tbody>
</table>

Before Chrome 58
## High-Level Numbers

<table>
<thead>
<tr>
<th>Crawl Dates</th>
<th>% Websites with sockets</th>
<th>% Sockets with A&amp;A Initiators</th>
<th>% Sockets with A&amp;A Receivers</th>
<th>#Unique A&amp;A Initiators</th>
<th>#Unique A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 02-05, 2017</td>
<td>2.1</td>
<td>60.6</td>
<td>73.7</td>
<td>75</td>
<td>16</td>
</tr>
<tr>
<td>Apr 11-16, 2017</td>
<td>2.4</td>
<td>61.3</td>
<td>74.6</td>
<td>63</td>
<td>18</td>
</tr>
<tr>
<td>May 07-12, 2017</td>
<td>1.6</td>
<td>60.2</td>
<td>69.7</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>Oct 12-16, 2017</td>
<td>2.5</td>
<td>63.4</td>
<td>63.7</td>
<td>23</td>
<td>18</td>
</tr>
</tbody>
</table>

Before Chrome 58

After Chrome 58
High-Level Numbers

<table>
<thead>
<tr>
<th>Crawl Dates</th>
<th>%Websites with sockets</th>
<th>% Sockets with A&amp;A Initiators</th>
<th>% Sockets with A&amp;A Receivers</th>
<th>#Unique A&amp;A Initiators</th>
<th>#Unique A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 02-05, 2017</td>
<td>2.1</td>
<td>60.6</td>
<td>73.7</td>
<td>75</td>
<td>16</td>
</tr>
<tr>
<td>Apr 11-16, 2017</td>
<td>2.4</td>
<td>61.3</td>
<td>74.6</td>
<td>63</td>
<td>18</td>
</tr>
<tr>
<td>May 07-12, 2017</td>
<td>1.6</td>
<td>60.2</td>
<td>69.7</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>Oct 12-16, 2017</td>
<td>2.5</td>
<td>63.4</td>
<td>63.7</td>
<td>23</td>
<td>18</td>
</tr>
</tbody>
</table>

- ~2% websites use web sockets.
## High-Level Numbers

<table>
<thead>
<tr>
<th>Crawl Dates</th>
<th>% Websites with sockets</th>
<th>% Sockets with A&amp;A Initiators</th>
<th>% Sockets with A&amp;A Receivers</th>
<th>#Unique A&amp;A Initiators</th>
<th>#Unique A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 02-05, 2017</td>
<td>2.1</td>
<td>60.6</td>
<td>73.7</td>
<td>75</td>
<td>16</td>
</tr>
<tr>
<td>Apr 11-16, 2017</td>
<td>2.4</td>
<td>61.3</td>
<td>74.6</td>
<td>63</td>
<td>18</td>
</tr>
<tr>
<td>May 07-12, 2017</td>
<td>1.6</td>
<td>60.2</td>
<td>69.7</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>Oct 12-16, 2017</td>
<td>2.5</td>
<td>63.4</td>
<td>63.7</td>
<td>23</td>
<td>18</td>
</tr>
</tbody>
</table>

- ~2% websites use web sockets.
- ~61% sockets are initiated by A&A domains

A&A = Advertising and Analytics
e.g. DoubleClick, Criteo, Adnxs
## High-Level Numbers

<table>
<thead>
<tr>
<th>Crawl Dates</th>
<th>%Websites with sockets</th>
<th>% Sockets with A&amp;A Initiators</th>
<th>% Sockets with A&amp;A Receivers</th>
<th>#Unique A&amp;A Initiators</th>
<th>#Unique A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 02-05, 2017</td>
<td>2.1</td>
<td>60.6</td>
<td>73.7</td>
<td>75</td>
<td>16</td>
</tr>
<tr>
<td>Apr 11-16, 2017</td>
<td>2.4</td>
<td>61.3</td>
<td>74.6</td>
<td>63</td>
<td>18</td>
</tr>
<tr>
<td>May 07-12, 2017</td>
<td>1.6</td>
<td>60.2</td>
<td>69.7</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>Oct 12-16, 2017</td>
<td>2.5</td>
<td>63.4</td>
<td>63.7</td>
<td>23</td>
<td>18</td>
</tr>
</tbody>
</table>

- ~2% websites use web sockets.
- ~61% sockets are initiated by A&A domains.
- ~71% sockets contact an A&A domain.

A&A = Advertising and Analytics
  e.g. DoubleClick, Criteo, Adnxs
## High-Level Numbers

<table>
<thead>
<tr>
<th>Crawl Dates</th>
<th>% Websites with sockets</th>
<th>% Sockets with A&amp;A Initiators</th>
<th>% Sockets with A&amp;A Receivers</th>
<th># Unique A&amp;A Initiators</th>
<th># Unique A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 02-05, 2017</td>
<td>2.1</td>
<td>60.6</td>
<td>73.7</td>
<td>75</td>
<td>16</td>
</tr>
<tr>
<td>Apr 11-16, 2017</td>
<td>2.4</td>
<td>61.3</td>
<td>74.6</td>
<td>63</td>
<td>18</td>
</tr>
<tr>
<td>May 07-12, 2017</td>
<td>1.6</td>
<td>60.2</td>
<td>69.7</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>Oct 12-16, 2017</td>
<td>2.5</td>
<td>63.4</td>
<td>63.7</td>
<td>23</td>
<td>18</td>
</tr>
</tbody>
</table>

- ~2% websites use web sockets.
- ~61% sockets are initiated by A&A domains
- ~71% sockets contact an A&A domain
- # Initiators drop after Chrome 58 release.

A&A = Advertising and Analytics
e.g. DoubleClick, Criteo, Adnxs
### High-Level Numbers

<table>
<thead>
<tr>
<th>Crawl Dates</th>
<th>%Websites with sockets</th>
<th>% Sockets with A&amp;A Initiators</th>
<th>% Sockets with A&amp;A Receivers</th>
<th>#Unique A&amp;A Initiators</th>
<th>#Unique A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 02-05, 2017</td>
<td>2.1</td>
<td>60.6</td>
<td>73.7</td>
<td>75</td>
<td>16</td>
</tr>
<tr>
<td>Apr 11-16, 2017</td>
<td>2.4</td>
<td>61.3</td>
<td>74.6</td>
<td>63</td>
<td>18</td>
</tr>
<tr>
<td>May 07-12, 2017</td>
<td>1.6</td>
<td>60.2</td>
<td>69.7</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>Oct 12-16, 2017</td>
<td>2.5</td>
<td>63.4</td>
<td>63.7</td>
<td>23</td>
<td>18</td>
</tr>
</tbody>
</table>

- ~2% websites use web sockets.
- ~61% sockets are initiated by A&A domains.
- ~71% sockets contact an A&A domain.
- # Initiators drop after Chrome 58 release.
- Small but persistent A&A receivers.

A&A = Advertising and Analytics
e.g. DoubleClick, Criteo, Adnxs
Initiators and Receivers
Initiators and Receivers

Initiator: JavaScript

Receiver
Initiators and Receivers

Initiator (JavaScript) \(\xrightarrow{ws/s}\) Receiver
Initiators and Receivers

Initiator: JavaScript

Receiver

ws/s
## Initiators and Receivers

**Initiator**

- **JavaScript**

**ws/s**

**Receiver**

### Top A&A Initiators

<table>
<thead>
<tr>
<th>A&amp;A Initiator</th>
<th>#A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>facebook</td>
<td>11</td>
</tr>
<tr>
<td>google</td>
<td>11</td>
</tr>
<tr>
<td>doubleclick</td>
<td>9</td>
</tr>
<tr>
<td>youtube</td>
<td>8</td>
</tr>
<tr>
<td>addthis</td>
<td>8</td>
</tr>
<tr>
<td>hotjar</td>
<td>7</td>
</tr>
<tr>
<td>googlesyndication</td>
<td>6</td>
</tr>
<tr>
<td>twitter</td>
<td>5</td>
</tr>
<tr>
<td>sharethis</td>
<td>4</td>
</tr>
<tr>
<td>adnxs</td>
<td>3</td>
</tr>
</tbody>
</table>
Initiators and Receivers

Initiator JavaScript ws/s Receiver

Top A&A Initiators

<table>
<thead>
<tr>
<th>A&amp;A Initiator</th>
<th>#A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>facebook</td>
<td>11</td>
</tr>
<tr>
<td>google</td>
<td>11</td>
</tr>
<tr>
<td>doubleclick</td>
<td>9</td>
</tr>
<tr>
<td>youtube</td>
<td>8</td>
</tr>
<tr>
<td>addthis</td>
<td>8</td>
</tr>
<tr>
<td>hotjar</td>
<td>7</td>
</tr>
<tr>
<td>googlesyndication</td>
<td>6</td>
</tr>
<tr>
<td>twitter</td>
<td>5</td>
</tr>
<tr>
<td>sharethis</td>
<td>4</td>
</tr>
<tr>
<td>adnxs</td>
<td>3</td>
</tr>
</tbody>
</table>
## Initiators and Receivers

### Top A&A Initiators

<table>
<thead>
<tr>
<th>A&amp;A Initiator</th>
<th>#A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>facebook</td>
<td>11</td>
</tr>
<tr>
<td>google</td>
<td>11</td>
</tr>
<tr>
<td>doubleclick</td>
<td>9</td>
</tr>
<tr>
<td>youtube</td>
<td>8</td>
</tr>
<tr>
<td>addthis</td>
<td>8</td>
</tr>
<tr>
<td>hotjar</td>
<td>7</td>
</tr>
<tr>
<td>googlesyndication</td>
<td>6</td>
</tr>
<tr>
<td>twitter</td>
<td>5</td>
</tr>
<tr>
<td>sharethis</td>
<td>4</td>
</tr>
<tr>
<td>adnxs</td>
<td>3</td>
</tr>
</tbody>
</table>

### Top A&A Receivers

<table>
<thead>
<tr>
<th>A&amp;A Receiver</th>
<th>#A&amp;A Initiators</th>
</tr>
</thead>
<tbody>
<tr>
<td>realtime</td>
<td>27</td>
</tr>
<tr>
<td>33across</td>
<td>19</td>
</tr>
<tr>
<td>intercom</td>
<td>16</td>
</tr>
<tr>
<td>disqus</td>
<td>13</td>
</tr>
<tr>
<td>zopim</td>
<td>12</td>
</tr>
<tr>
<td>hotjar</td>
<td>11</td>
</tr>
<tr>
<td>feedjit</td>
<td>10</td>
</tr>
<tr>
<td>lockerdome</td>
<td>8</td>
</tr>
<tr>
<td>inspectlet</td>
<td>6</td>
</tr>
<tr>
<td>smartsupp</td>
<td>4</td>
</tr>
</tbody>
</table>
Initiators and Receivers

Top A&A Initiators

<table>
<thead>
<tr>
<th>A&amp;A Initiator</th>
<th>#A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>facebook</td>
<td>11</td>
</tr>
<tr>
<td>google</td>
<td>11</td>
</tr>
<tr>
<td>doubleclick</td>
<td>9</td>
</tr>
<tr>
<td>youtube</td>
<td>8</td>
</tr>
<tr>
<td>addthis</td>
<td>8</td>
</tr>
<tr>
<td>hotjar</td>
<td>7</td>
</tr>
<tr>
<td>googlesyndication</td>
<td>6</td>
</tr>
<tr>
<td>twitter</td>
<td>5</td>
</tr>
<tr>
<td>sharethis</td>
<td>4</td>
</tr>
<tr>
<td>adnxs</td>
<td>3</td>
</tr>
</tbody>
</table>

Top A&A Receivers

<table>
<thead>
<tr>
<th>A&amp;A Receiver</th>
<th>#A&amp;A Initiators</th>
</tr>
</thead>
<tbody>
<tr>
<td>realtime</td>
<td>27</td>
</tr>
<tr>
<td>33across</td>
<td>19</td>
</tr>
<tr>
<td>intercom</td>
<td>16</td>
</tr>
<tr>
<td>disqus</td>
<td>13</td>
</tr>
<tr>
<td>zopim</td>
<td>12</td>
</tr>
<tr>
<td>hotjar</td>
<td>11</td>
</tr>
<tr>
<td>feedjit</td>
<td>10</td>
</tr>
<tr>
<td>lockerdome</td>
<td>8</td>
</tr>
<tr>
<td>inspectlet</td>
<td>6</td>
</tr>
<tr>
<td>smartsupp</td>
<td>4</td>
</tr>
</tbody>
</table>

Disqus provides comment board services.
Initiators and Receivers

**Top A&A Initiators**

<table>
<thead>
<tr>
<th>A&amp;A Initiator</th>
<th>#A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>facebook</td>
<td>11</td>
</tr>
<tr>
<td>google</td>
<td>11</td>
</tr>
<tr>
<td>doubleclick</td>
<td>9</td>
</tr>
<tr>
<td>youtube</td>
<td>8</td>
</tr>
<tr>
<td>addthis</td>
<td>8</td>
</tr>
<tr>
<td>hotjar</td>
<td>7</td>
</tr>
<tr>
<td>googlesyndication</td>
<td>6</td>
</tr>
<tr>
<td>twitter</td>
<td>5</td>
</tr>
<tr>
<td>sharethis</td>
<td>4</td>
</tr>
<tr>
<td>adnxs</td>
<td>3</td>
</tr>
</tbody>
</table>

**Top A&A Receivers**

<table>
<thead>
<tr>
<th>A&amp;A Receiver</th>
<th>#A&amp;A Initiators</th>
</tr>
</thead>
<tbody>
<tr>
<td>realtime</td>
<td>27</td>
</tr>
<tr>
<td>33across</td>
<td>19</td>
</tr>
<tr>
<td>intercom</td>
<td>16</td>
</tr>
<tr>
<td>disqus</td>
<td>13</td>
</tr>
<tr>
<td>zopim</td>
<td>12</td>
</tr>
<tr>
<td>hotjar</td>
<td>11</td>
</tr>
<tr>
<td>feedjit</td>
<td>10</td>
</tr>
<tr>
<td>lockerdome</td>
<td>8</td>
</tr>
<tr>
<td>inspectlet</td>
<td>6</td>
</tr>
<tr>
<td>smartsupp</td>
<td>4</td>
</tr>
</tbody>
</table>

- **Disqus** provides comment board services.
- **Zopim, Intercom, Smartsupp** provide live chat services.
# Initiators and Receivers

## Top A&A Initiators

<table>
<thead>
<tr>
<th>A&amp;A Initiator</th>
<th>#A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>facebook</td>
<td>11</td>
</tr>
<tr>
<td>google</td>
<td>11</td>
</tr>
<tr>
<td>doubleclick</td>
<td>9</td>
</tr>
<tr>
<td>youtube</td>
<td>8</td>
</tr>
<tr>
<td>addthis</td>
<td>8</td>
</tr>
<tr>
<td>hotjar</td>
<td>7</td>
</tr>
<tr>
<td>googlesyndication</td>
<td>6</td>
</tr>
<tr>
<td>twitter</td>
<td>5</td>
</tr>
<tr>
<td>sharethis</td>
<td>4</td>
</tr>
<tr>
<td>adnxs</td>
<td>3</td>
</tr>
</tbody>
</table>

## Top A&A Receivers

<table>
<thead>
<tr>
<th>A&amp;A Receiver</th>
<th>#A&amp;A Initiators</th>
</tr>
</thead>
<tbody>
<tr>
<td>realtime</td>
<td>27</td>
</tr>
<tr>
<td>33across</td>
<td>19</td>
</tr>
<tr>
<td>intercom</td>
<td>16</td>
</tr>
<tr>
<td>discus</td>
<td>13</td>
</tr>
<tr>
<td>zopim</td>
<td>12</td>
</tr>
<tr>
<td>hotjar</td>
<td>11</td>
</tr>
<tr>
<td>feedjit</td>
<td>10</td>
</tr>
<tr>
<td>lockerdome</td>
<td>8</td>
</tr>
<tr>
<td>inspectlet</td>
<td>6</td>
</tr>
<tr>
<td>smartsupp</td>
<td>4</td>
</tr>
</tbody>
</table>

- **Disqus** provides comment board services.
- **Zopim, Intercom, Smartsupp** provide live chat services.
- **33across & Lockerdome** are advertising platforms.

---

**Diagram:**

- Initiator (JavaScript) to Receiver (ws/s) connection diagram.
## Initiators and Receivers

![Diagram showing Initiator and Receiver with JavaScript as the Initiator and Receiver as the Receiver]

### Top A&A Initiators

<table>
<thead>
<tr>
<th>A&amp;A Initiator</th>
<th>#A&amp;A Receivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>facebook</td>
<td>11</td>
</tr>
<tr>
<td>google</td>
<td>11</td>
</tr>
<tr>
<td>doubleclick</td>
<td>9</td>
</tr>
<tr>
<td>youtube</td>
<td>8</td>
</tr>
<tr>
<td>addthis</td>
<td>8</td>
</tr>
<tr>
<td>hotjar</td>
<td>7</td>
</tr>
<tr>
<td>googlesyndication</td>
<td>6</td>
</tr>
<tr>
<td>twitter</td>
<td>5</td>
</tr>
<tr>
<td>sharethis</td>
<td>4</td>
</tr>
<tr>
<td>adnxs</td>
<td>3</td>
</tr>
</tbody>
</table>

### Top A&A Receivers

<table>
<thead>
<tr>
<th>A&amp;A Receiver</th>
<th>#A&amp;A Initiators</th>
</tr>
</thead>
<tbody>
<tr>
<td>realtime</td>
<td>27</td>
</tr>
<tr>
<td>33across</td>
<td>19</td>
</tr>
<tr>
<td>intercom</td>
<td>16</td>
</tr>
<tr>
<td>disqus</td>
<td>13</td>
</tr>
<tr>
<td>zopim</td>
<td>12</td>
</tr>
<tr>
<td>hotjar</td>
<td><strong>11</strong></td>
</tr>
<tr>
<td>feedjit</td>
<td>10</td>
</tr>
<tr>
<td>lockerdome</td>
<td>8</td>
</tr>
<tr>
<td>inspectlet</td>
<td><strong>6</strong></td>
</tr>
<tr>
<td>smartsupp</td>
<td>4</td>
</tr>
</tbody>
</table>

- **Disqus** provides comment board services.
- **Zopim, Intercom, Smartsupp** provide live chat services.
- **33across & Lockerdome** are advertising platforms.
- **Inspectlet & Hotjar** are session replay services.
Sent Items Over Web Sockets
Sent Items Over Web Sockets

% Requests

0 20 40 60 80

Cookie
IP
User IDs
Fingerprinting
Variables
DOM

WebSocket
HTTP/S
Stateful Identifiers like Cookie and User IDs

- Cookie
- IP
- User IDs
- Fingerprinting Variables
- DOM

Sent Items Over Web Sockets

% Requests

- WebSockets
- HTTP/S

- Stateful Identifiers like Cookie and User IDs
• Stateful Identifiers like Cookie and User IDs

• Fingerprinting data in ~3.4% WebSockets. 97% is 33across
Sent Items Over Web Sockets

- Stateful Identifiers like Cookie and User IDs
- Fingerprinting data in ~3.4% WebSockets. 97% is 33across
- ~1.6% WebSockets sends the entire DOM to Hotjar, LuckyOrange, TruConversion
Received Items Over Web Sockets
Received Items Over Web Sockets

- HTML
- JSON
- JavaScript
- Images

% Responses

- WebSockets
- HTTP/S
Received Items Over Web Sockets

- HTML
- JSON
- JavaScript
- Images

% Responses

WebSockets
HTTP/S
Received Items Over Web Sockets

- HTML
- JSON
- JavaScript
- Images

% Responses

- WebSockets
- HTTP/S
Received Items Over Web Sockets

<table>
<thead>
<tr>
<th>Item</th>
<th>WebSockets</th>
<th>HTTP/S</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTML</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JSON</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JavaScript</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Images</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

% Responses

Ads served from Lockerdome
Summary

• ~67% of socket connections are initiated or received by A&A domains.

• Major companies like Google, Facebook, Addthis adopted WebSockets. Abandoned after Chrome 58 was released.

• The culprits:
  
  • **33across** was harvesting fingerprinting data.
  
  • DOM exfiltration by **HotJar, LuckyOrange, TruConversion**
  
  • **Lockerdome** downloaded URLs to serve ads.

• We need to keep up with the current practices of A&A companies.
Summary

• ~67% of socket connections are initiated or received by A&A domains.

• Major companies like Google, Facebook, Addthis adopted WebSockets. Abandoned after Chrome 58 was released.

• The culprits:
  
  • **33across** was harvesting fingerprinting data.

  • DOM exfiltration by **HotJar, LuckyOrange, TruConversion**

  • **Lockerdome** downloaded URLs to serve ads.

• We need to keep up with the current practices of A&A companies.

Questions?
ahmad@ccs.neu.edu
Backup Slides
Inclusion Chain

DOM Tree

```
<html>
  <body>
    <script src="tracker/script.js" />
    <img src="tracker/img.jpg">
    <script src="ads/script.js">
    <iframe src="frame.html">
      <html>
        <body>
          <script src="script_12.js">
            <img src="img_a.jpg">
        </body>
      </html>
    </iframe>
  </body>
</html>
```

Inclusion Tree

```
pub/index.html
  tracker/script.js
    tracker/img.jpg
      ads/script.js
        ads/frame.html
          ads/script_12.js
            adnet/data.ws
              Source code for ads/script_12.js
                let ws =
                  new WebSocket("ws://adnet/data.ws", …);
                  ws.onopen = function (e) {ws.send("…");}