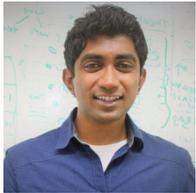


Adblocking and Counter-Blocking



Rishab
Nithyanand



Sheharbano
Khattak



Mobin
Javed



Narseo Vallina-
Rodriguez



Marjan
Falahrastegar



Julia E.
Powles



Emiliano De
Cristofaro



Hamed
Haddadi



Steven
Murdoch

Now... our sponsors



OPEN TECHNOLOGY FUND



THE ROYAL
SOCIETY



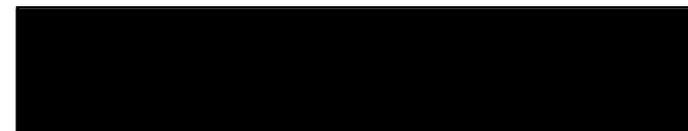
EPSRC

Engineering and Physical Sciences
Research Council

xerox



Skipping slide in 15
seconds



Or... Turn off your adblocker to proceed



In this paper, we ...

- Understand the current state of the content provider vs. adblocker arms race

Quick Facts:

- Ad blocking estimated to cost publishers nearly \$22 billion
- There are now 198 million active adblock users around the world
- Ad blocking grew by 41% globally in the last 12 months
- US ad blocking grew by 48% to reach 45 million active users
- UK ad blocking grew by 82% to reach 12 million active users

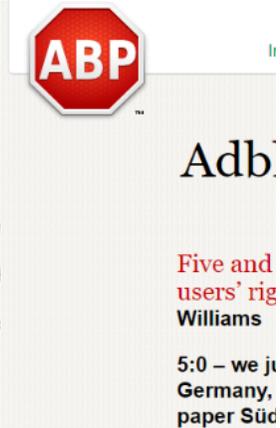


DETECT
EXPLAIN
ASK
LIFT or LIMIT

- **Detect** ad blocking, in order to
- **Explain** the value exchange to
- **Ask** for changed behavior in exchange
- **Lift** restrictions or **Limit** access



IAB Tech Lab Publisher Adblocking Primer



Eyo will



And immediately face malvertising scares
Security and privacy activists fight back

Popular websites deploy anti-adblockers

Adobe and PageFair 2015 at

August '15

March '16

March '16

April '16



In this paper, we ...

- Understand the current state of the content provider vs. adblocker arms race
 - How widespread are anti-adblockers?
 - Anti-adblockers: Adblocker blockers and detectors.
 - What sites are employing them?
 - Who are the most popular vendors of anti-adblockers?
 - How do the popular anti-adblockers work?
 - Can adblockers block anti-adblockers?
 - Or, can adblockers block adblocker blockers?

How widespread are anti-adblockers?

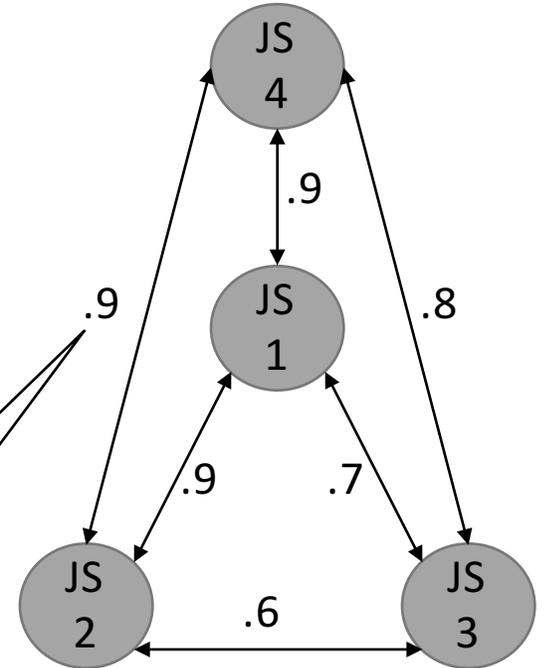
- Crawl and download all JSes in the Alexa T-5K.
 - Over 200K JS in Alexa T-5K.
 - Analyzing (functionality of) each one is not feasible.
- Observation: Most sites use the same JSes.
 - Bought from vendors, part of a library, etc.
- Cluster similar JSes and analyze each cluster.

Clustering similar JSes

- Formulate as the Clique Finding problem.
- Find all cliques with edge weight $\geq x$ (say, .80).

- JS 1, JS 2, and JS 4 are likely to have identical functionality and source for sufficiently high “ x ”.

- Gives a lower-bound of shared JSes. Fails with differently obfuscated code and composed JSes.



Similarity of code (TF-IDF)

White-space insensitive

Position insensitive

Resistant to noise

Analyzing clusters of JSes

- Characteristics (script vendor and functionality) of each clique are identified by analyzing:
 - Sources of scripts.
 - External resources utilized by the scripts.
 - (TF-IDF) Keywords associated with the script.
 - **Manually.**

How widespread are anti-adblockers?

- Total number of cliques with size ≥ 5 : 1882
 - 1882 JSes shared by 5+ websites.
- Total number of tracker JSes: 456 of 1882
 - 456 JSes used for tracking shared by 2741 websites.
- **Total number of anti-adblocker JSes: 22 of 1882**
 - **22 JSes used for detecting adblockers shared by 335 websites.**
 - This is a lower bound.

What sites are employing anti-adblockers?

- McAfee's URL categories over the 335 identified websites.

%	Category	%	Category
19.5%	General News	2.5%	Pornography
9.3%	Blogs/Wiki	2.5%	Forum/Bulletin Boards
8.5%	Entertainment	2.2%	Technical/Business Forums
4.3%	Internet Services	2.2%	Potential Illegal Software
3.7%	Sports	2.0%	Online Shopping
3.7%	Games	1.7%	Portal Sites
3.2%	Travel	1.7%	Humor/Comics
3.2%	Education/Reference	1.2%	Social Networking
2.7%	Business	1.2%	Provocative Attire
2.5%	Software/Hardware	1.2%	Marketing/Merchandising

Who are the popular anti-adblocking vendors?

Domain	Description	#Sites
pagefair.com	Anti-adblocking	20
googleadservices.com	Ads	61
googlesyndication.com	Ads	13
taboola.com	Ads	36
outbrain.com	Ads	10
ensighten.com	Ads	6
hotjar.com	Analytics	9
doublepimp.com	Pornography	8
tacdn.com	Travel	8
cloudflare.com	CDN	50
cloudfront.net	CDN	6
yimg.com	Content/Ads	108

How do anti-adblockers work?

- Key idea: Bait objects.
- Example 1: Using a bait DIV element

```
var test = document.createElement('div'); //Create a new DIV element
test.innerHTML = '&nbsp;';
test.className = 'adsbox'; //Set bait

document.body.appendChild(test); //Add element to document body

if (test.offsetHeight === 0) //Check element height
{
    document.body.classList.add('adblock');
}
```

How do anti-adblockers work?

- Key idea: Bait objects.
- Example 2: Using bait JSes

```
// ads.js  
var foo = False
```

```
// main.js  
var foo = True  
// main.html  
if (foo is True)  
{  
    document.body.classList.add('adblock');  
}
```

Can adblockers block anti-adblockers?

Domain	Description	#Sites	ABP	Gh	PB
pagefair.com	Anti-adblocking	20	✓	✗	✓
googleadservices.com	Ads	61	✗	✗	✗
googlesyndication.com	Ads	13	✗	✗	✗
taboola.com	Ads	36	✗	✓	✓
outbrain.com	Ads	10	✗	✓	✓
ensighten.com	Ads	6	✗	✓	✗
hotjar.com	Analytics	9	✗	✗	✗
doublepimp.com	Pornography	8	✗	✓	✗
tacdn.com	Travel	8	✗	✗	✗
cloudflare.com	CDN	50	✗	✗	✓
cloudfront.net	CDN	6	✗	✗	✗
yting.com	Content/Ads	108	✗	✗	✗

TL; DR

- Anti-adblockers are employed by **at least** 6.7% of the Alexa T-5K.
- Anti-adblockers are mostly employed by “News” websites.
- Anti-adblockers are being supplied mostly by ad vendors.
 - PageFair report should be taken with a grain of salt (conflict of interest).
- The arms race has entered the next phase.
 - Escalation seems likely unless a compromise is found (e.g., Eyeo AAI).

Upcoming work

- Automated categorization of JSes.
- Longitudinal measurements.