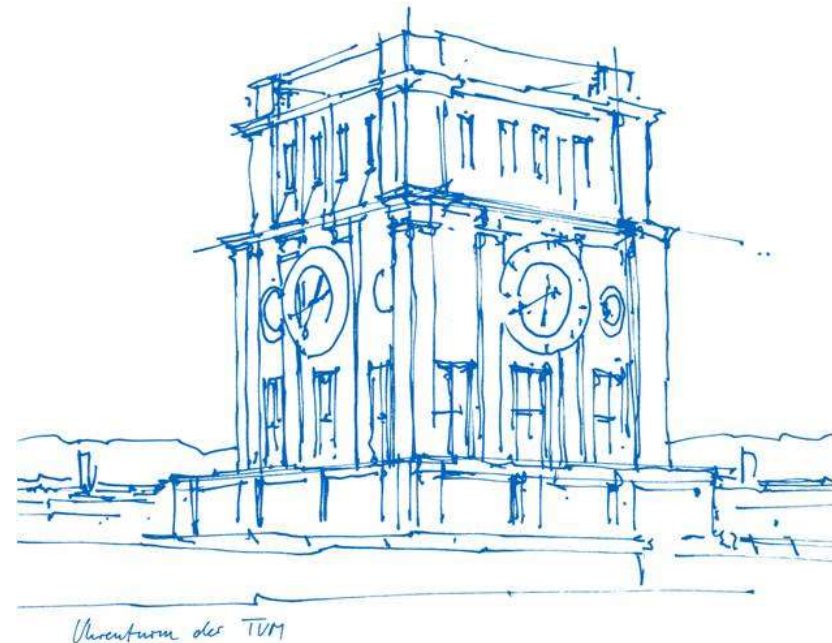# Blocking the Blocker – Studying the effects of Anti Ad-blocking

Rohit Gupta & Rohit Panda

Technische Universität München

Fakultät für Informatik

München, 7. August 2018

Introduction

- Online advertisements drives the economy of the World Wide Web.

- Promotional content (ads) are now strategically placed on webpages to attract attention of the users.

- Users are installing Ad-blocker tools citing interference with user experience, privacy and protection against malware as the top reasons.

- Publishers are employing Anti Ad-blockers, scripts and notification mechanisms that ask users to disable any existing installed Ad-blocker tools.

Roadmap

- In our paper, we study various aspects of Anti Ad-blockers including:

  - We study the usage of Ad-blocker scripts and their mechanisms.

  - We collect and analyse data from top websites using Alexa Web Ranking.

  - We discuss the methodology used in detecting Anti Ad-blockers and present our results.

  - We discuss the economic impact of Anti Ad-blockers and how GDPR affects Anti Ad-blockers.

  - We discuss the legality and ethical aspects of using Anti Ad-blockers.

  - We discuss alternatives to Anti Ad-blocking.

Ad-blocker popularity in Germany

- PageFair Adblock penetration   vs   BVDW* and OVK[+] page impressions.

Adblock penetration per online capita, %
PageFair

	Desktop	Mobile	Overall
U. States	18%	1%	18%
China	1%	13%	13%
U. Kingdom	16%	1%	16%
Japan	3%	–	3%
Germany	29%	1%	29%
Australia	20%	2%	20%
Canada	24%	–	25%
France	11%	1%	11%
Brazil	6%	1%	6%
South Korea	4%	–	4%
Russia	6%	3%	6%
Italy	17%	1%	17%
Netherlands	17%	2%	17%
Spain	19%	–	19%
Mexico	8%	–	9%
Sweden	27%	–	27%
Indonesia	8%	58%	58%
Denmark	25%	2%	25%
India	1%	28%	28%
Argentina	14%	–	14%
Finland	23%	–	23%

Share of Desktop Page Impressions in Germany on Which Display Ads Were Blocked, Q2 2015-Q3 2016
% of total

Q2 2015: 21.5%
Q3 2015: 21.2%
Q4 2015: 21.5%
Q1 2016: 20.1%
Q2 2016: 19.4%
Q3 2016: 19.1%
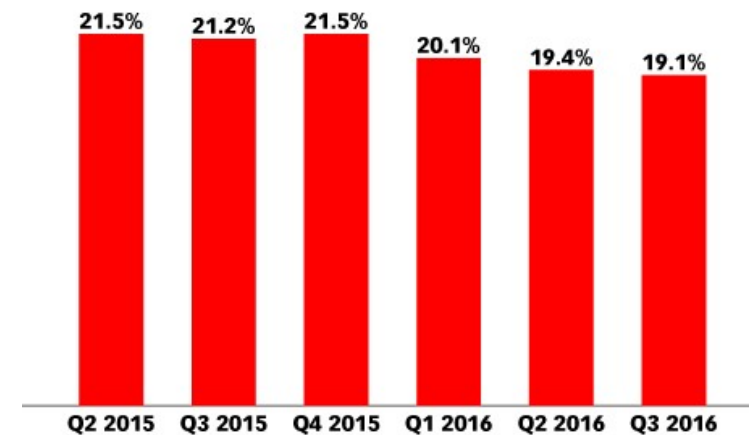
Note: based on figures reported by OVK members; read as delivery of ads was blocked on 19.1% of page impressions in Q3 2016
Source: Bundesverband Digitale Wirtschaft (BVDW) and Online-Vermarkterkreis (OVK), "Zentrale Adblocker-Rate des OVK" as cited in company blog, Nov 16, 2016
220131   www.eMarketer.com

Source: PageFair and www.eMarketer.com

How Ad-blockers work

- Code showing how Forbes implements ads on their website:

```
1  j = document.createElement("div"),
2  j.setAttribute("class", r.join(" ")),
3  document.body.appendChild(j), "none" === window.getComputedStyle(j).display
4  && (this.removeWelcomeCookies(), this.adblock_state = "on")
5  }
6  }, this.removeWelcomeCookies = function() {
7          var b = function() {
8              s = !0, window.advBidxc && advBidxc.adBlockDetected
9               window.uabfunc && window.uabfunc.detected
10             (a.remove("welcomeAd", {
11                 path: "/",
12                 domain: ".forbes.com"
13             }), a.remove("dailyWelcomeCookie", {
14                 path: "/",
15                 domain: ".forbes.com"
16             }), a.put("re_ab", 1, {
17                 path: "/",
18                 domain: ".forbes.com"
19             }))
20         };
```

How Ad-blockers work

- Code showing how Forbes implements ad-blocker detection on their website:

```
1  <script>
2     try {
3           fbs_settings.blocked_classes = ["sidebar_ADBOX","textlinkads",
4        "advertisement-banner","externalAdComponent","ad_label2a",
5        "adseparator","content-advertisment","review_ad1","loop-ad",
6        "BottomGoogleAds","ad-160-160","aopsadvert","sidebar_advertising",
7        "yan-sponsored","advertBox","ad-parent-hockey","block-maniad",
8        "advertising-block","google_ad_wide","txtAd","rightSideSponsor",
9        "adwolf-holder","view-display-id-ads_all","ad_links",
10       "advertisement-tag","widget_ad_rotator","ad-caption",
11       "ads-bottom-block","gemini-ad","rightColAdBox"];}
12    catch(err) {
13          fbs_settings.blocked_classes = null;
14    }
15 </script>
16 <script type="text/javascript">setTimeout(function() {
17    document.getElementById('main-content').className = '';
18    }, 1000);
19    try {
20    performance.mark('blocking_scripts_start');
21    } catch (e) {}
22 </script>
23 <script src="//i.forbesimg.com/forbes/scripts/b4b56484.vendor.js"></script>
24 <script src="//i.forbesimg.com/forbes/scripts/0d9abc37.scripts.js"></script>
25 <div id="css-js-dynamic"><span class="dynamic-css">false</span>
26    <span class="dynamic-js"></span>
27 </div>
28 <div ng-if="!$root.is_mobile" id="teconsent"></div>
29 <script>$(window).on("touchstart", function(e){});</script>
30 <script type="text/javascript">
31    try {
32          performance.mark('blocking_scripts_end');
33          } catch (e) {}
34 </script>
```

Ad-blocker extensions

- µBlock Origin for Chrome and Firefox
- AdBlock for Chrome
- Adblock Plus for Firefox, Chrome, Opera and Safari
- AdBlock Pro for Chrome
- Adguard for Chrome and Firefox
- AdRemover for Chrome
- Ghostery for Chrome, Firefox, Opera, Safari, Internet Explorer, Android and iPhone iOS
- Simply Block Ads! for Chrome
- SuperBlock AdBlocker for Chrome
- µ Adblock for Firefox
- uMatrix for Firefox, Chrome and Opera

Data Collection

- We collected top websites from Amazon Alexa Website Rankings

- Publicly only top 50 websites for each category is available, we subscribed to a 7-day free trial to collect the following data sets:
  - Top 500 websites in Germany
  - Top 500 websites each in Germany, Austria and Switzerland (DACH region) and created a unique subset of these websites.
  - Top 500 websites in the category: News

Sample HTTP request to pull data from AWIS:

```
https://awis.amazonaws.com/api?Action=CategoryListings&Count=20
  &Descriptions=True&Path=Top%2FNews
  &Recursive=False&ResponseGroup=Listings
  &SortBy=Popularity&Start=1
```

Detecting Anti Ad-blockers: Motivation

- Our papers of reference are:

    - Haris et al. "Detecting Anti Ad-blockers in the wild"

    - Rishab et al. "Adblocking and Counter-Blocking: A Slice of the Arms Race"

- Their papers used Alexa top 100K websites which bias towards a US and China specific website list.

- Our focus is on Germany, DACH and a specific category News.

Detecting Anti Ad-blockers: Overview

- Anti Ad-blocker scripts detect and change HTML or DOM contents by:

  - adding extra HTML elements,

  - change in style of existing HTML elements,

  - changes in textual context.

- As described by Haris et al. in their paper. these changes could be:

  - *Nodal Changes*: Adding extra DOM elements

  - *Style Changes*: Modifying style and appearance of DOM elements

  - *Textual Changes*: Changes in textual content

  - *Structural Changes*: Such as URL redirections

Detecting Anti Ad-blockers: Methodology

- We make use of Selenium Chrome Webdriver to launch two separate instances of Google Chrome for A/B testing.

- Instance A is launched with Adblock Plus extension and instance B is without any extension.

- Instances are compared w.r.t Node Features (div, h1, h2, h3, img, table, p, iframe and text nodes), Textual Features (#lines, #words, #text) and Structural Features (URL redirection).

- These features are compared using a Python script and BeautifulSoup library is used for recording the difference and screen scraping.

- These features are stored in a .csv file.

Detecting Anti Ad-blockers: Demo
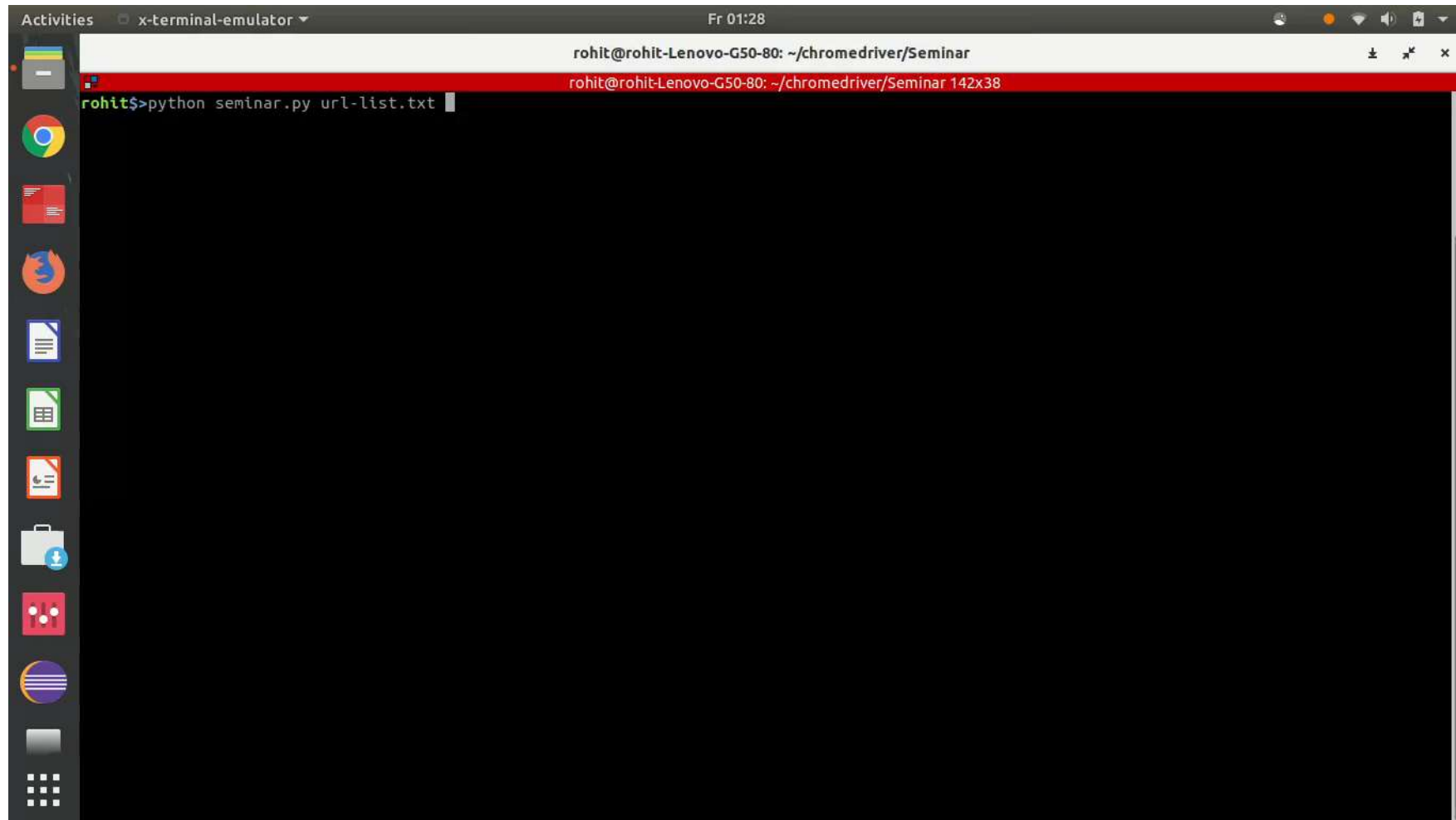
Detecting Anti Ad-blockers: Model Training

- We use three Machine Learning classifiers: J48 Decision Tree, Random Forest and Naive Bayes to identify websites that employ anti Ad-blockers.

- We collected positive samples (websites that employ Anti Ad-blockers) and negative samples (websites that do not employ Anti Ad-blockers) from sources such as anti-adblock-killer list and Alexa top-100 websites.

- Information gain ratio is used for determining which features extracted are most important for learning and prediction.

Detecting Anti Ad-blockers: Feature Analysis

- Following are the features ranked based on the Information Gain:

Information Gain	Feature
25.65%	lines
20.66%	p
20.60%	a
20.44%	div
19.2%	words
18.05%	tags
16.12%	img
11.76%	h1
11.05%	keyword
11.05%	h3
9.15%	iframe
8.67%	h2
4.87%	table
0.78%	url change

Detecting Anti Ad-blockers: Classifier Evaluation

- Decision Tree visualization for Anti Ad-blockers (using Weka toolkit):
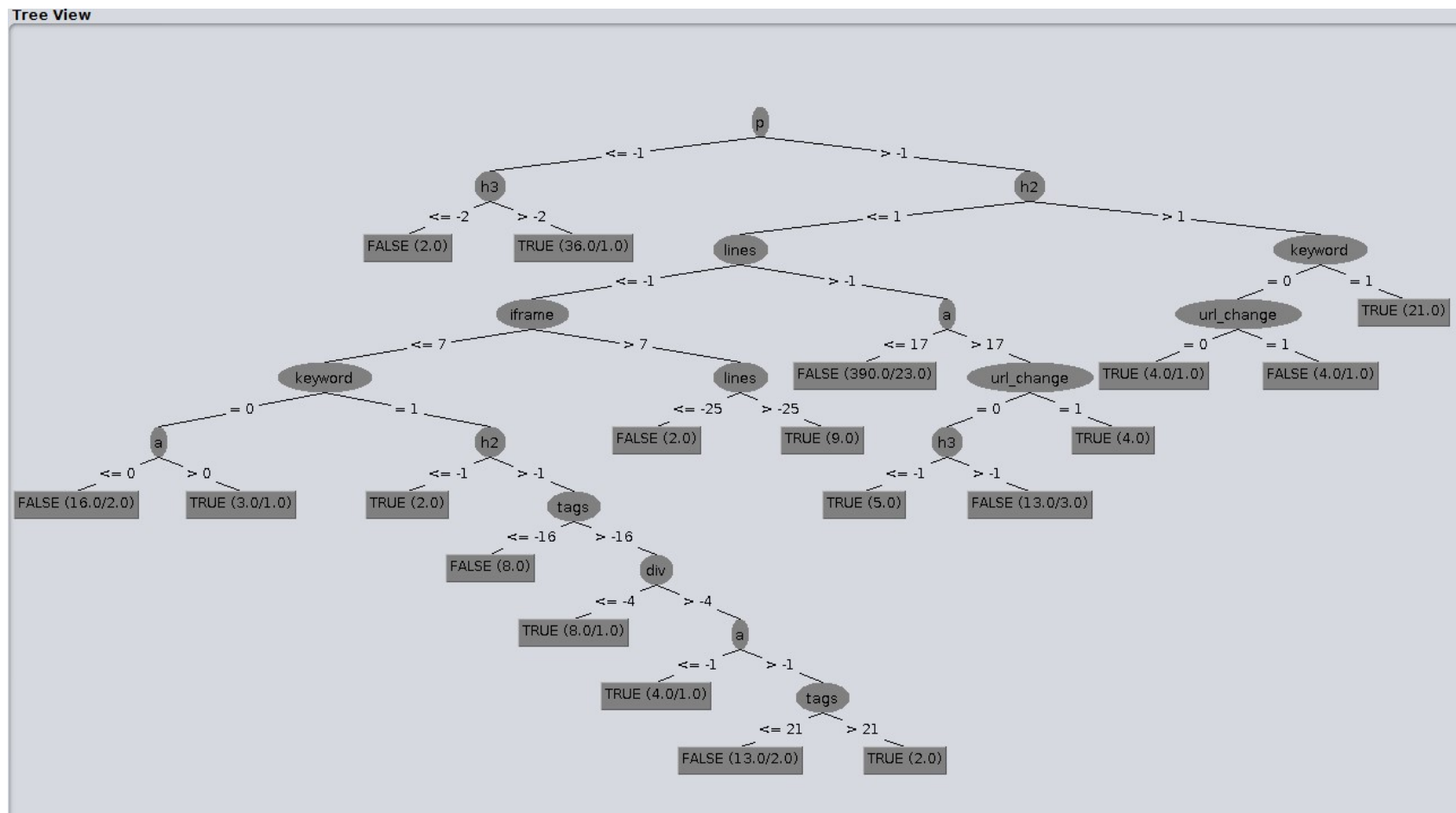
Detecting Anti Ad-blockers: Classifier Evaluation

- Detailed Accuracy by Class for Random Forest classifier:

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	1.000	0.024	0.993	1.000	0.996	0.984	0.999	1.000	FALSE
	0.976	0.000	1.000	0.976	0.988	0.984	0.999	0.998	TRUE
Weighted Avg.	0.995	0.019	0.995	0.995	0.994	0.984	0.999	0.999	

- Confusion Matrix for Random Forest classifier :

a	b	<- classified as
422	0	a = FALSE
3	121	b = TRUE

- Also done for Bayes Naive and J48 Decision Tree using Weka toolkit.

Datasets and Results: Germany

- From the list of top 500 websites in Germany using Alexa Website Rankings, we were able to successfully create a test set of 418 websites:

	Naive Bayes Classifier	J48 Classifier	Random Forest Classifier
Predicted	4	24	26
TP	3	8	8
FP	1	16	18
Precision	0.75	0.333	0.308

- 8 out of the 418 websites (1.9%) top visited German websites employ Ad-block detection.

Datasets and Results: DACH region

- From the list of top 1500 websites in Germany, Austria and Switzerland using Alexa Website Rankings, we were able to successfully create a test set of 809 unique websites:

	Naive Bayes Classifier	J48 Classifier	Random Forest Classifier
Predicted	10	40	41
TP	3	11	10
FP	7	29	31
Precision	0.3	0.275	0.244

- 11 out of the 809 websites (1.4%) top visited German websites employ Ad-block detection.

- The low precision is attributed to a lot of websites deploying data protection (GDPR) information on the main page or are allowed under Acceptable Ads programme.

Datasets and Results: News category

- From the list of top 500 websites in the News Category using Alexa Website Rankings, we were able to successfully create a test set of 357 websites:

	Naive Bayes Classifier	J48 Classifier	Random Forest Classifier
Predicted	7	18	16
TP	6	10	11
FP	1	8	5
Precision	0.857	0.556	0.688

- 11 out of the 357 websites (3.1%) top visited news websites employ Adblock detection.

- The high precision for News websites is synonymous with the results presented by our reference papers by Haris et al. and Rishab et al.

Ad-block Detection Responses: "CIA" measure

- We verified the positive samples for Ad-block detector responses.

- We define a new measure called the "CIA" measure for such responses:

  - Cost Model (C): We describe this type as a cost or monetization model.

  - Invisibility (I): We describe this type as not allowing users to view content until Ad-blockers are disabled.

  - Availability (A): We describe this type as a conservative approach, allowing users to view content but asking them politely to disable Ad-blockers.

Ad-block Detection Responses: 'C'IA example
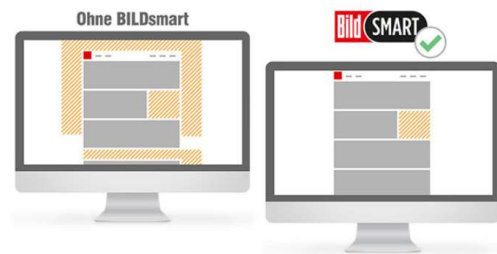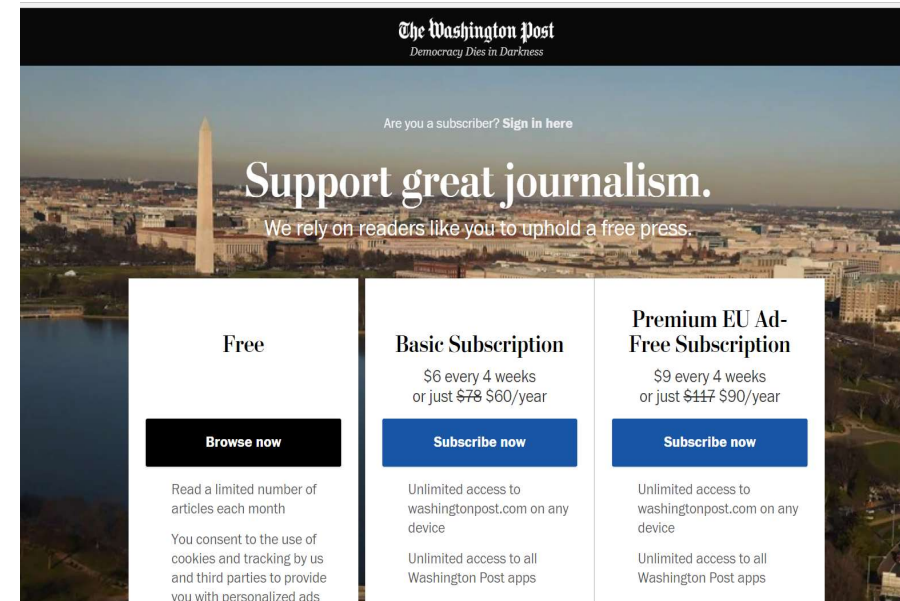
Source: www.bild.de, www.washingtonpost.com and www.pornhub.com

Ad-block Detection Responses: C'I'A example

Source: www.sport1.de

Ad-block Detection Responses: CI'A' example

Source: www.t3n.de

# Datasets and Results: Geographical Comparison

- Our reference paper by Haris et al. focussed on Alexa top-100K websites that includes majority of U.S.A. and China websites.

- They found 686 websites in Alexa top-100K (6.7%) deploy Ad-blocker detection.

- In our analysis we found that 1.9% websites (out of 418 German websites) and 1.4% (out of 809 websites in DACH region) employ some kind of Ad-blocker detection.

Limitations

- We remove all anti Ad-block filters in the Adblock Plus' filter list. This may not be the default configuration.

- Our methodology focusses only on homepage of such websites.

- We consider only HTML, DOM and textual changes to websites.

- Our focus is on Adblock Plus due to its popularity.

- We do not consider style features and cosine similarity used by our reference papers.

Economical Impact

- U.S. leads the market share with $58.13b revenue in 2017 followed by the U.K. with $11.72b and China with $10.81b.
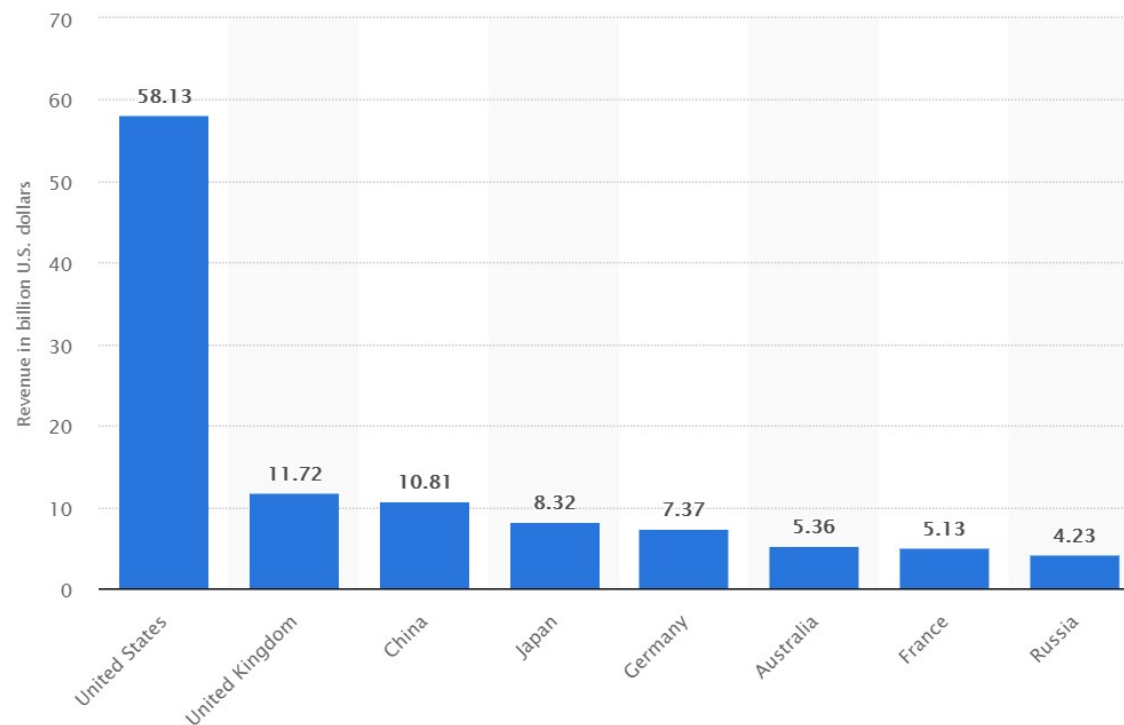
Source: www.statista.com

Anti Ad-blockers: Legality in Germany

- There have been many legal battles fought till date, e.g., between Adblock Plus company Eyeo and Sueddeutsche Zeitung, Pro-Sieben-Sat.1, and IP Deutschland. However, German courts have always ruled in favour of Ad-blockers.

- The Munich higher regional court did not classify Eyeo's business model as "forbidden aggressive advertising".

- It is legal!

Anti Ad-blockers: Ethical aspect

- Ads are the biggest sources of revenue for publishers on the world wide web.

- Many websites have come up with a subscription model.

- Some publishers, part of the Acceptable Ads programme, may not reveal all details regarding their deal with Ad-blocker companies.

- Here is where the "CIA" measure can be used to classify if the business model is ethical or not.

Anti Ad-blockers: Impact of GDPR

- The General Data Protection Regulation (GDPR) does not allow companies to store or use user's personal data unless explicitly agreed by the user.

- However, detecting Ad-blockers work mostly on the concept of identifying HTML and DOM elements which are completely unrelated to the users (still complying with GDPR).

- Websites that do not fully understand the risks of data protection are only providing full text content or country redirection links to an ad-free EU version of their websites.

Anti Ad-blocker Alternatives: Acceptable Ads

- Acceptable ads programme is provided by Adblock Plus for providing ads that do not disrupt or distort web content for users. The following criteria must be followed for an ad to be an Acceptable Ad:

  - *Placement*

  - *Distinction*

  - *Size*

- Unacceptable ads include animated ads, autoplay, expanding ads, pop-ups and pop-unders, rich media ads.

Anti Ad-blocker Alternatives: Whitelists

- This is analogous to anti-viruses that allow certain softwares to run on the system.

- Adblock Plus charges high generating revenue sources for adding their websites to their whitelists while they do not charge for low income revenue sources.

YES! Anti Ad-blocker killers also exists!

- Anti Ad-block killers tricks sites that use Anti Ad-blocker detection into thinking the user is not using an Ad-blocker.

- Anti Ad-block killers such as AAK are composed of a user script AakScript written in JavaScript with a default filter list similar to that of AdBlock and AdBlock Plus.

- Ad-blockers still continue to run as normal using script managers like Greasemonkey or Tampermonkey.

Conclusion

- 1.4 – 1.9% websites use Ad-block detection scripts for a specific country or region, whereas 3.1% websites use these scripts for a specific category such as News.

- For us, Bayes Naive classifier generates the highest precision in contrast to our reference paper by Haris et al.

- German courts have ruled in favour of Ad-blocker extensions till date and GDPR as such does not directly affect the working of such scripts.

References

- Haris et. al. Detecting Anti Ad-blockers in the wild. Proceedings on privacy enhancing technologies (2015)

- Rishab et. al. Ad-blocking and counter blocking: A slice of the arms race (2016)

- Garimella et al. Ad-blocking: A study on performance, privacy and counter-measures (2017)

- Amazon Alexa Website Rankings: https://www.alexa.com/topsites

- Digital advertising in Europe - Statistics and Facts (2017)
  https://www.statista.com/topics/3983/digital-advertising-in-europe/

Thank you!