

The Ad Wars: Retrospective Measurement and Analysis of Anti-Adblock Filter Lists

Umar Iqbal
The University of Iowa

Problem Statement

Online ads have several problems: they intrude user privacy, they are intrusive, they spread malware, and they have performance overheads

To get rid of all of the problems with ads people use adblockers



However, adblockers impact publishers revenue and publishers fight back with anti-adblockers

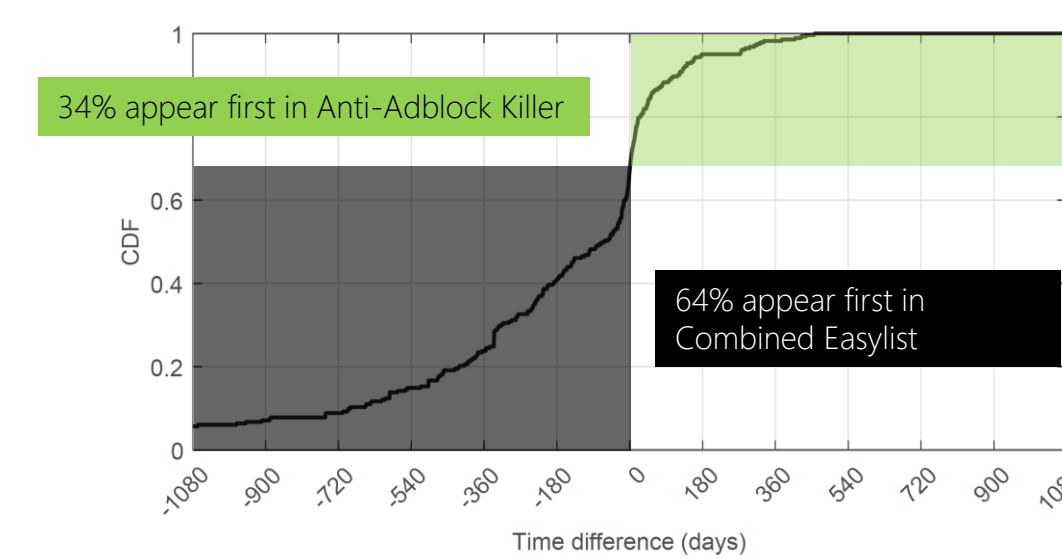
Solution: anti-adblocking filter lists!



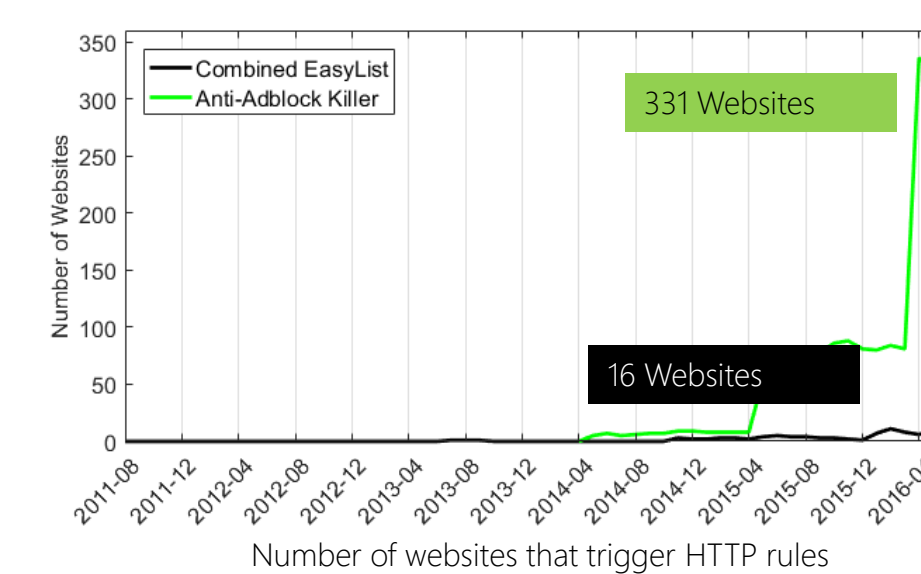
We analyze the effectiveness of anti-adblock filter lists and propose a machine learning approach to automate them

Retrospective Coverage Analysis of Anti-Adblock Filter Lists

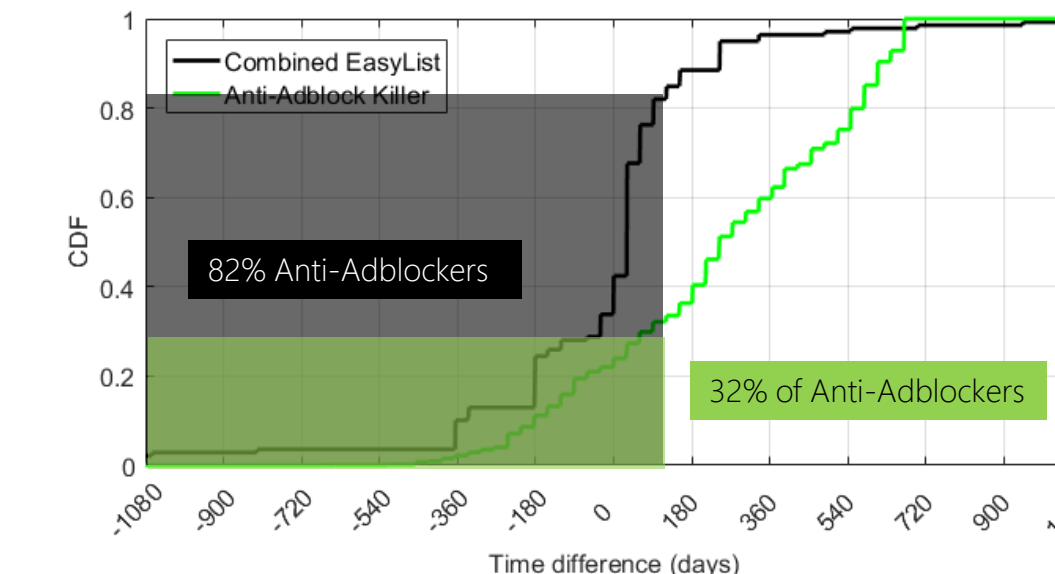
Popular filter lists: Anti-Adblock Killer and Combined EasyList (EasyList + Warning Removal List)



Combined EasyList is More Prompt in Adding New Rules



Anti-Adblock Killer Filter List Has Better Coverage



Combined EasyList Updates More Frequently

Results & Evaluation

10-fold cross-validation
Evaluation on Alexa top 100K (92.5% TPR)

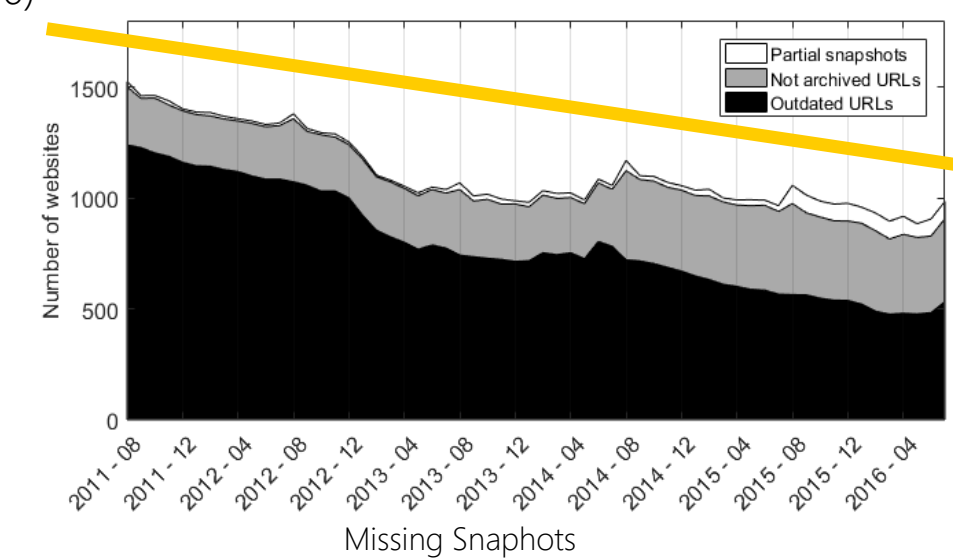
Feature Set	Classifier	Number of Features	TP rate (%)	FP rate (%)
all	AdaBoost + SVM	10K	99.6	3.9
literal	AdaBoost + SVM	10K	99.6	3.9
keyword	AdaBoost + SVM	1K	99.7	3.2

Complement manual analysis

- Periodic crawl to expedite manual process
- Substantial reduction of manual effort

The Wayback Machine

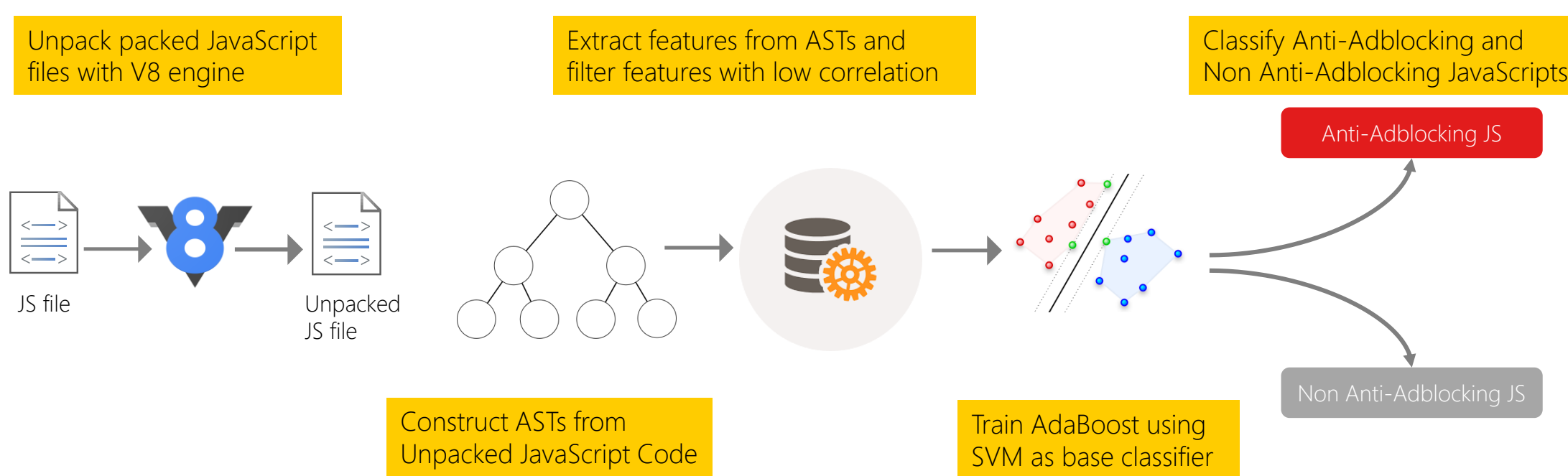
Archives webpages and all their resources (279 billion webpages)
Crawled data of Alexa top 5K websites of 5 years (2011 - 2016)



The Wayback Machine Has Improved Over Time!

Anti-Adblock Scripts Detection With Machine Learning

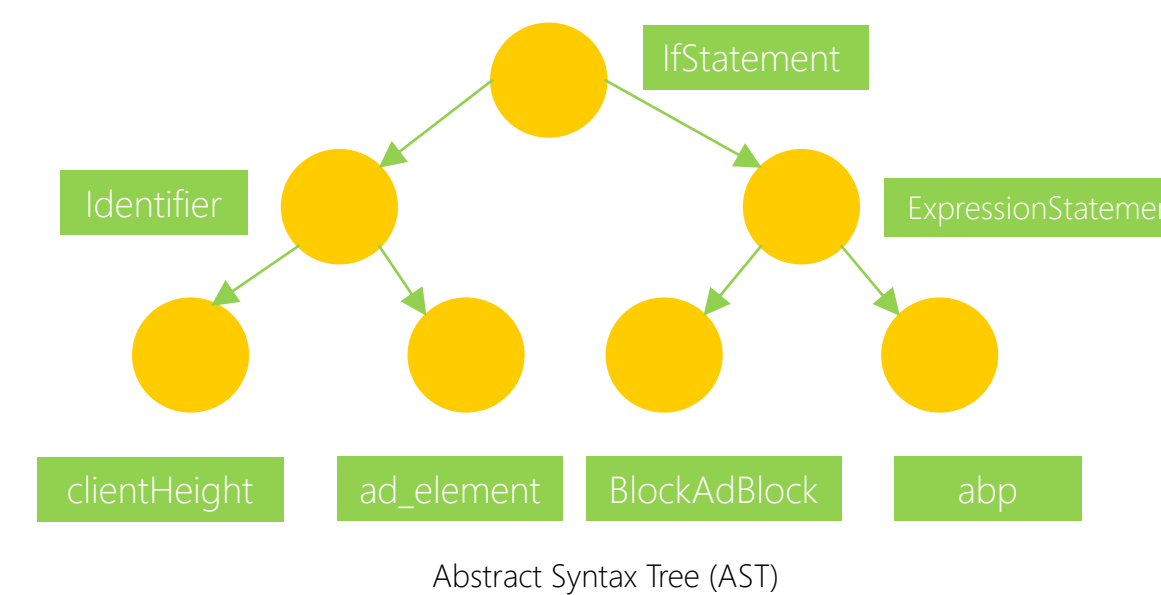
Anti-Adblocking code from 3rd party vendors, code has structural similarities
Static analysis to capture code structure



Feature Extraction

Features (context : text)

- All (AssignmentExpression:BlockAdBlock)
- Literal (Literal:abp)
- Keyword (Identifier:clientHeight)



Key Takeaways

Retrospective measurement of anti-adblocking filter lists
Lightweight machine learning approach to complement filter lists
The Wayback Machine's potential for retrospective web measurements

More details in our IMC'17 paper:
The Ad Wars: Retrospective Measurement and Analysis of Anti-Adblock Filter Lists
Umar Iqbal, Zubair Shafiq, Zhiyun Qian
ACM Internet Measurement Conference (IMC), 2017.