Review of Perceptual Ad Blocking Technologies

Yuhui Lyu  
ykl5501@psu.edu

**ABSTRACT**

Traditional ad blockers rely on popular public filter lists which are maintained and updated by crowdsourcing. Some computer scientists propose a new type of adblocking technique called perceptual ad blocking. They believe that this new perceptual ad blocker is almost impossible for web publishers and ad networks to defeat. The perceptual ad blocker will be the next-generation ad blocking system and may end the arms race between users and ad networks. In this report, I will review techniques behind different perceptual ad blockers and discuss several cases and applications of ad blockers using perceptual signals as the reason for figuring out which ads to hide on pages.

**1. INTRODUCTION**

Advertising services are the main income sources of web publishers like Google and Facebook. However, people are gradually unable to tolerate more and more advertisements, which seriously affect the user experience. The concerns about security problems like third party tracking is also a reason why people need adblocker. Traditional ad blockers heavily rely on popular public filter lists which are maintained and updated by crowdsourcing. The users of filter lists, such as EasyList, provide feedback and report errors to editors in the community. The editors will double-check feedback or errors and then update the blocking rule and filter list manually.[1] Currently, most ad blockers collect and maintain a large dataset of ad metadata including page markup and URLs.[2] In order to counter and evade the detection of ad blockers, web publishers and ad net-works apply changes in their website source codes to obscure ad disclosures. An arms race between ad blockers and web publishers begins: ad blockers add more detection and blocking rules to fight the changes of source codes of pages; publishers deploy more changes to evade the detection. Arvind Narayanan, a computer scientist at Princeton University, proposed and developed a new concept of ad blocking technology called "Perceptual Ad Blocker" which may terminate this arms race. The perceptual ad blocker will act and perceive online ads like a human by using machine learning and computer vision techniques. [3] Inspired by this idea, some popular ad blockers also develop (or plan to develop) their perceptual ad blocking system. This article will conclude concepts of perceptual ad blockers and different techniques behind them.

**2. RELATED WORK**

Most current ad blockers rely on manual, crowdsourced maintenance, which is hard to catch up with the pace of the dynamic evolution of the advertisement landscape. Some authors proposed new automated techniques, especially machine learning techniques, to automatically complement ad blocking filter lists with less manual effort. [4]

Most ad blockers are designed to filter the elements in pages. When we talk about perceptual ad blockers, we often assume that their blocking mechanism is based on the visual signals. Adblock Radio, however, is a perceptual ad blocker that will detect and block the audio ads in radio-based on audio signals. Adblock Radio consists of two ad detectors. The first detector is to analyze the texture and pattern of audio. Since ads are distinct from actual contents and their patterns are universal, the first detector could filter most ads using a machine learning algorithm with a little maintenance. The second detector uses the Shazam-like algorithm and a crowdsourced database to filter the ad which is ignored by the first detector.

There is an implicit agreement between web publishers who provides free online s contents and services and subscribers. With the emergence of ad blockers, it becomes a huge threat to the publisher's revenues from advertising. Therefore, more and more web publishers are willing to design anti-adblocking systems to detect the existence of ad blockers and their users. [5] This phenomenon declares a new arms race between ad-blockers and anti-adblockers.

**3. METHODS**

**Optical character recognition (OCR).** Many traditional ad blockers detect keywords (ads disclosures) like “ads” and “Sponsored” in pages to recognize the advertisements. In order to obfuscate them, web publishers like Facebook use increasingly complex HTML obfuscation. [2] For example, a “Sponsored” tag could be divided into several parts: a “Sp” tag, a “on” tag, a “so” tag, and a “red” tag. Humans could still understand the meaning of the tag, but it is hard for ad blockers. Perceptual ad blockers using OCR could evade this obfuscation since they perceive ads like humans.

Some web publishers convert plaintexts of ad closure to images to evade the detection of traditional ad blockers. Therefore, perceptual ad blockers could apply OCR to distinguish the text in rendered images on web pages.

**Image matching and object detection.** The premise of concepts of the perceptual ad blocker is that humans could recognize ads by glancing at pages and perceptual ad blockers could simulate this process. Federal Trade Commission (FTC) regulates the federal advertising standards and stipulates that advertising must be truthful, non-deceptive, and cannot be unfair. [6] AdChoices, which is an industrial standard icon, can be used as a sign of online behavior advertising. Many advertisers and web publishers adopt this standard and add AdChoices icon into their online ads. In turn, perceptual ad blockers utilize different computer vision techniques such as convolutional neural networks, such as R-CNN, Fast R-CNN, and YOLO, to detect the AdChoices icon to recognize the ads on web pages.

A picture containing drawing

Description automatically generated

Figure 1: The AdChoices icon.

**Template matching.** Some ad blockers like Ad-Highlighter apply algorithms of template matching to detect the ad disclosure logo. They use average hashing, scale-invariant feature transform (SIFT), and other feature detection algorithms to compare the features of the template to those of images on a web page.

A screenshot of a cell phone

Description automatically generated

Figure 2: An example that the perceptual ad blocker detects ad disclosure and highlight it.

**Click simulator.** Some perceptual ad blockers intimate human behaviors and analyze behavior features. They will simulate a mouse click operation like a human open a new link, and check whether the browser would be redirected to a new page which contains advertisements or ad disclosures. The blocker will check the destination URL and its contents to recognize the ads.

**Case: PERCIVAL.** PERCIVAL is a light-weight perceptual ad blocker using a deep learning network. It works on the browser's image rendering pipeline and can capture all processed and rendered images during page execution. The development team of PERCIVAL chooses a small CNN network SqueezeNet to classify each intercepted image as either an or not an ad. The biggest difference between PERCIVAL and other perceptual ad blockers using CNN, such as Sentinel and Ad Highlighter, is that PERCIVAL receives images directly from the image decoding pipeline rather than taking rendered contents of a page. [7]

**4. EVALUATION**

Most perceptual ad blocks are still in the concept or development stage. Therefore, it is hard to evaluate their performance by myself. I conclude the evaluation results and data from two papers. The first paper is about a perceptual ad blocker developed by Grant Storey’s team. The second paper is about Percival, a new in-browser perceptual ad blocker using deep learning. I also create a small demo to detect the ad disclosure on images collect from multiple websites.

**The ad blocker of Grant Storey’s team.** The team evaluated the effectiveness of their ad blocker extension. They collected 50 ads on Facebook and the extension detected all 50 ads accurately. In the next 6 months, the team did not notice any false positive or negative errors.

The team evaluated the extension on 100 websites randomly. They manually labeled 212 AdChoices ads and the ad blocker correctly identified most of them, with only 3 FP and 4 FN errors. [3]

**A screenshot of a cell phone

Description automatically generated** Figure 3: Evaluation result of identifying advertisements contains the AdChoices icon.

**Percival.** The team first compared the performance and accuracy of their product with a famous and popular ad blocking list EasyList. They used a crawler to obtain images from news websites in Alexa top 500 websites and manually labeled images. PERCIVAL could identify ads with accuracy 96.76%, precision 97.76%, and recall 95.72%. [7]

A picture containing table

Description automatically generated

Figure 4: Performance of Percival against Easylist

The team tested Percival on Facebook. They used Facebook for 35 days and found that Percival could detect ads with accuracy 92%, precision 78.4%, and recall 70%.

A screenshot of a cell phone

Description automatically generated

Figure 5: Performance of Percival on Facebook

They also applied Percival on Google Image and Percival performed differently based on the query results. The team tested Percival on both lowly ad-biased and highly ad-biased keywords and Percival blocked more images when the team queried contents with highly ad-biased.

A screenshot of a computer

Description automatically generated

Figure 6: Percival performance on Google images for searching “advertisement”.

**Demo.** I designed a small demo which detects the ad disclosure in images. I captured images of ads from several news websites. The demo uses OCR to detect the text “AdChoices”.

A close up of a person

Description automatically generated

Figure 7: A demo using OCR to detect the text “AdChoices”.

I also trained an Alexnet to detect the AdChoices logo to simulate the detection operation of other perceptual ad blockers.

A picture containing person, sitting, computer, sign

Description automatically generated

Figure 8: A demo using the trained network to detect the AdChoices logo.

The demo does not have a stable performance. Sometimes, the logo or the text of AdChoices is too small to detect for the demo. I also do not have too much experience in deep learning and computer vision. I will try to optimize it in the future, but the ideas and concepts of the demo are on the right track.

**5. Discussion and Conclusion**

Traditional ad blockers use crowdsourcing filter list to detect the target ads. The maintenance and update of a filter list require lots of laborious efforts. The arms race between traditional ad blockers and web publishers and advertisers will never end. The perceptual ad blocker will be a perfect solution for these problems since we assume that a perceptual ad blocker could perceive ads like a human and all ads could be identified by human eyes.

It seems that a perceptual ad blocker is perfect and undefeatable. However, it may not be the case. Now, most of the perceptual ad blockers are still in the planning or designing stages. Customers cannot find a practical and useful perceptual ad blocker. Moreover, researchers from Stanford University also reveal the potential vulnerabilities of perceptual ad blockers and conduct several attacks that could easily bypass or spoof the detection of perceptual ad blockers. [2] Some people believe that the arms race between ad blockers and publishers will end one day. I think this is impossible since nothing is perfect, and we could always find an approach to break through the obstacle. Although perceptual ad blocking is not perfect, this technology still has great potential.

**References**

[1] M. Alrizah, X. Xing, S. Zhu, and G. Wang, “Errors, misunderstandings, and attacks: Analyzing the crowdsourcing process of ad-blocking systems,” 2019.

[2] F. Tramer, P. Dupre, G. Rusak, G. Pellegrino, and D. Boneh, “Adversarial: Perceptual adblocking meets adversarial machine learning,” 2019.

[3] G. Storey, D. Reisman, J. Mayer, and A. Narayanan, “The future of ad blocking: An analytical framework and new techniques,” 2017.

[4] D. Gugelmann, M. Happe, B. Ager, and V. Lenders, “An automated approach for complementing ad blockers’ blacklists,” 2015.

[5] M. H. Mughees, Z. Qian, and Z. Shafiq, “Detecting anti ad-blockers in the wild,” 2017.

[6] Federal Trade Commission, “Advertising faq’s: A guide for small business.” [Online]. Available: <https://www.ftc.gov/tips-advice/business-center/guidance/advertising-faqs-guide-small-business>.

[7] Z. A. Din, P. Tigas, S. T. King, and B. Livshits, “Percival: Making in-browser perceptual ad blocking practical with deep learning.” [Online]. Available: <https://www.usenix.org/system/files/atc20-din.pdf>.