# Your Online Interests – Pwned!
# A Pollution Attack Against Targeted Advertising

Wei Meng
Georgia Institute of Technology
wei@gatech.edu

Xinyu Xing
Georgia Institute of Technology
xxing8@gatech.edu

Anmol Sheth
Technicolor
anmolsheth@gmail.com

Udi Weinsberg
Technicolor
udi.weinsberg@gmail.com

Wenke Lee
Georgia Institute of Technology
wenke@cc.gatech.edu

## ABSTRACT

We present a new ad fraud mechanism that enables publishers to increase their ad revenue by deceiving the ad exchange and advertisers to target higher paying ads at users visiting the publisher's site. Our attack is based on polluting users' online interest profile by issuing requests to content not explicitly requested by the user, such that it influences the ad selection process. We address several challenges involved in setting up the attack for the two most commonly used ad targeting mechanisms – re-marketing and behavioral targeting. We validate the attack for one of the largest ad exchanges and empirically measure the monetary gains of the publisher by emulating the attack using web traces of 619 real users. Our results show that the attack is *effective* in biasing ads towards the desired higher-paying advertisers; the polluter can influence up to 74% and 12% of the total ad impressions for re-marketing and behavioral pollution, respectively. The attack is *robust* to diverse browsing patterns and online interests of users. Finally, the attack is *lucrative* and on average the attack can increase revenue of fraudulent publishers by as much as 33%.

## Categories and Subject Descriptors

J.0 [**Computer Applications**]: General

## General Terms

Security

## Keywords

Online Advertising; Ad Fraud; Profile Pollution; Ad Measurement

## 1. INTRODUCTION

Online targeted advertising is one of the primary approaches used to monetize free online services and applications available to users. Recently, there has been a concerted effort to increase the relevance of ads targeted at users by tailoring the ads to their stated or inferred interests. Studies have shown that ads targeted based on a user's online interests have a 40% higher chance in leading to a financial conversion over non-targeted ads [24]. Consequently, the average price online advertisers and marketers pay for these targeted ads is 2.6 times higher than non-targeted ads [15].

The revenue model for online targeted advertising can be described by the function of three primary entities: *advertisers*, *publishers* and *ad exchange platforms*. Ad exchange platforms (*e.g.*, DoubleClick) facilitate the buying and selling of ads between the advertiser and publisher. Publishers register their website with the ad exchange and host ad slots. Advertisers set up campaigns by describing their target audience, *e.g.*, specifying user demographics and interests, along with a maximum cost they are willing to pay for ad impressions or clicks made by their target audience. The ad exchange runs an online auction based on the bid values received from all the competing advertisers, and delivers the winning ad to the user visiting the publisher page. The revenue generated from this transaction is shared between the publisher (which typically receives 68% [3] of the revenue) and the ad exchange.

As is evident from the above description, there are two main factors that impact the publisher's revenue. The first is the number of users visiting the publisher webpage which in turn impacts the number of ad impressions or clicks served by the publisher. The second is the cost that advertisers are willing to pay to have their ads targeted at users visiting the publisher page.

In this paper we present a new ad fraud mechanism that enables publishers to increase their ad revenue by exploiting the role played by the user's online interest profile in the ad selection process. Our attack exploits the fact that advertisers mainly set up campaigns to target users with specific online interests and are willing to pay higher for such users. Since the user's interest profile is inferred based on the webpages a user visits, it is vulnerable to exploits that use Cross-Site Request Forgery (CSRF) [18], clickjacking [27] or cross-site scripting (XSS) [32] that can pollute users' profiles by generating camouflaged requests to webpages not explicitly visited by them. A fraudulent publisher can use these exploits to pollute the profiles of users visiting the publisher's website to mislead advertisers and the ad exchange to deliver more lucrative ads to these users, and thereby increase the publisher's ad revenue.

While the above described attack seems intuitive, it is not trivial to design and launch the attack such that it is practical, effective, and lucrative. To the best of our knowledge, our paper is the first to design and successfully deploy a pollution attack on the existing


*CCS'14,* November 3–7, 2014, Scottsdale, Arizona, USA.

http://dx.doi.org/10.1145/2660267.2660273.

targeted advertising ecosystem. Achieving this requires addressing the following challenges which also form the main contributions of our work. First, the attack should not require any explicit cooperation from the ad exchange or advertisers, and should be effective for the two commonly used ad targeting mechanisms – behavioral targeting and re-marketing. Second, polluting user profiles should be effective even without explicit knowledge about external factors that impact ad revenue (campaign budgets, bid costs, publisher preferences and ad inventory, *etc.*). Third, it should be feasible to load the pollution content in a camouflaged manner such that it is not discernible by the users while deceiving the ad exchange and advertisers. Finally, the polluted user profile should result in biasing the ads targeted at the user towards the intended higher-paying advertisers.

To address the above described challenges, we set up and validate the attack against one of the largest ad exchanges, DoubleClick, and study the monetary value of the attack for live publisher webpages. Instead of polluting live traffic, we emulate user traffic to the publisher websites by replaying web traces collected from 619 real users from 264 distinct IP addresses and recording all ads delivered to these emulated users. This setup enables an end-to-end characterization of the different aspects of the attack under controlled settings that is otherwise not feasible. Our results show that the attack is *successful* and *effective* in deceiving DoubleClick to deliver higher-paying ads on the fraudulent publisher's website. Using our attack, the polluter can influence up to 74% and 12% of the total ad impressions for re-marketing and behavioral pollution, respectively. Finally, we show that the attack is *lucrative*, enabling the fraudulent publishers to increase their ad revenue on average by 33%.

The rest of this paper is structured as follows. We provide an overview of the ad targeting mechanisms in Section 2. In Section 3 we detail the profile pollution attack and discuss its deployment challenges. Section 4 describe the setup we use to validate the attack in a real-world deployment, and Section 5 details the validation results. We quantify the expected increase in revenue obtained from deploying the attack on real websites in Section 6. In Section 7 we discuss potential countermeasures that can help mitigate the attack. Finally, we discuss related work in Section 8 and conclude in Section 9.

# 2. AD TARGETING AND USER PROFILES

In this section we describe the ad targeting mechanisms available to advertisers [2] and discuss the critical role played by a user's online interest profile in the existing ad ecosystem.

## 2.1 Ad Targeting Mechanisms

**Contextual Targeting.** Contextual targeting involves matching the ad with the context of the page that it is displayed on (and ignores the visitor interest profile). The targeting is implicit and the user's online interests are largely ignored: a car insurance company will place ads on auto-related sites because it is assumed that visitors to the site are likely to own a car (or want to) and will need insurance.

**Re-Marketing.** Re-marketing is used by advertisers to target users who, in the past, have indicated a very specific interest in a particular product. For example, consider a user who visits a car insurance website, clicks on a link to get a quote, but leaves without buying the insurance offered. The insurance company (via the ad exchange) can then target this user with re-marketing ads, *e.g.*, showing insurance discounts. These ads will be delivered to the user on other websites, which may be completely unrelated to cars or insurance, to lure the user back to finish the purchase. Here, the advertiser targets a user by exploiting a very specific signal.

**Behavioral Targeting.** Behavioral targeting is used by advertisers that target users who have shown an interest in some categories (*e.g.*, cars or college football). This mechanism goes beyond the "single domain" aspect of re-marketing, and selects ads that might relate to the user's online interests as observed from her browsing patterns. This form of targeting often results in ads that may be unrelated with the page being viewed [29]. For example, with behavioral targeting, a user might be targeted with car insurance related ads (potentially from a company she did not visit online) on a website about Food & Nutrition simply because she visited multiple different car insurance related websites, and the ad exchange profiled her to be interested in car insurance.

## 2.2 User Profiles and Targeted Ads

Behavioral targeting and re-marketing make explicit use of the user's online interests that are profiled by the ad exchange and other third party trackers. This is achieved by installing third party JavaScript tracking code provided by the ad exchange on websites that users' browse. The tracking code extracts details about the page (*e.g.*, exact URL, `meta` tags about keywords, description, *etc.* [35]) and transmits this along with the user's cookie identifier. This information, along with other information that the ad exchange has about the website, is used to profile the user's interests and are offered to advertisers as targeting options. A user's interest profile for behavioral targeting is represented as a set of semantic categories, structured as a hierarchy (*e.g.*, Movies→Action Films→Superhero films). For re-marketing, the ad exchange simply maintains a list of users (cookie IDs) that visit a specific page on the advertiser's website.

As is evident, the user's interest profile forms an integral component of the ad selection process. Advertisers assign a monetary value using cost-per-click (CPC) or cost-per-mille (CPM) directly to the user's online interests and are willing to pay up to 2.6 times higher to target ads at users with a desired profile [15]. Moreover, as we show in Section 5, although a user may have online interests accumulated over a long time period, short term browsing activity can significantly impact the user's profile and consequently change the type of ads that a user receives. Our attack exploits this critical aspect and enables publishers to pollute user profiles towards ad categories that generate higher revenue. In the following section we provide an overview of the attack and present techniques for profile pollution that are specific to the ad ecosystem and the commonly used ad targeting mechanisms.

# 3. PROFILE POLLUTION ATTACK

The profile pollution attack (which we also refer to as a fraud mechanism) introduces a new entity in the ad ecosystem that we call *profile polluter*. Figure 1 shows the interaction of the profile polluter with the rest of the ad ecosystem (dashed lines). Specifically, the primary steps involved in a successful attack are:

1. The profile polluter identifies and downloads content in order to pollute user profiles.
2. A user visits the polluter page (which can be hosted at the publisher's website) and pollution content is loaded first in a camouflaged manner. (steps 1 and 1a in the figure).
3. This signals the ad exchange of a legitimate browsing event by the user, and the user's profile is impacted (step 1b).
4. When the user navigates to another page on the publisher's website, the ad exchange is deceived in using this modified profile in soliciting bids for ads (steps 2-5).

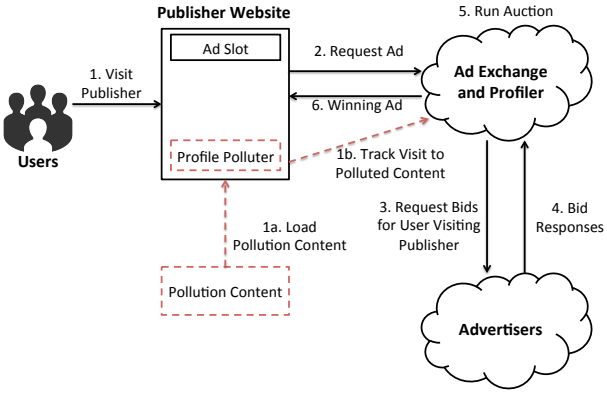



Figure 1: An overview of the profile pollution attack

5. The publisher's revenue increases if the winning ad is from an advertiser that bids higher to target the polluted user (step 6).

The attack focuses on polluting users to influence behavioral targeting and re-marketing ad campaigns, as they explicitly make use of the user's online interest profile. In order to simplify the description, we assume that the publisher also plays the role of the profile polluter. The readers should note under this assumption the attack could only impact ads of user's next visit, as the website content and pollution content are loaded by browser in parallel in each visit.

## 3.1 Identifying Pollution Content

There are two key requirements for selecting pollution content. First, content selected for pollution should alter the user profiles towards advertisers that bid *higher* to target users. Since information about advertising budgets and bid prices is typically not shared externally, the polluter can resort to aggregate revenue reports generated by the ad exchange. For example, data published by DoubleClick [26] lists the top three display advertising categories that generate the highest CPM as *Health*, *Business* and *Job & Education*. The profile polluter selects websites for the two pollution mechanisms from these categories. More importantly, the different categories and their associated revenue estimates enables the attacker to control the monetization of the attack and potentially go undetected by the ad exchange by not raising any anomalous revenue alerts. Second, the selected content needs to contain the tracking cookies and code used by the ad exchange to track and profile users. This deceives the ad exchange that the pollution content is a regular browsing activity, and consequently successfully alter the user's profile.

**Pollution Content for Re-Marketing.** A re-marketing campaign is set up by integrating a few lines of JavaScript code, *i.e.*, the *re-marketing script*, which is provided by the ad exchange, in the advertiser's website. The JavaScript code encodes the unique identifier of the advertiser and the associated re-marketing campaign. When a user visits the re-marketing enabled advertiser website, these identifiers along with the user's ad exchange cookie are transmitted to the ad exchange. This enables the ad exchange to tag the user and track her interactions on the advertiser's website. The tagged user is then easily re-identified later on other websites and is targeted with ads from the advertiser. Consequently, a user's past

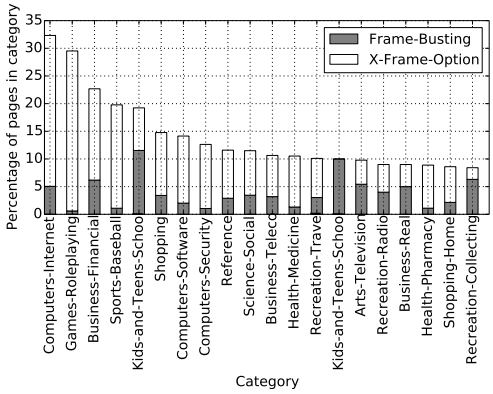



Figure 2: The percentage of websites that use frame-busting and X-Frame-Options techniques in the Alexa top 20 categories

browsing history and online interests do not impact re-marketing ads.

This script can be easily detected by parsing the HTML code of a webpage[1]. Thus, a simple approach to find content for re-marketing pollution is to parse webpages of advertisers belonging to high-paying categories and identify those that host re-marketing scripts.

**Pollution Content for Behavioral Targeting.** The approach of simply scanning websites in a directory service is not sufficient for finding content for behavioral pollution as the ad exchange categories may not match those of the directory service. Alternatively, the polluter can exploit the ad preference dashboards made available by large ad exchanges to build an offline map between webpages and the category label assigned to these webpage. Specifically, the polluter can impersonate a user with a blank profile (delete all cookies and create a fresh browser profile), browse pages from a specific category and record the corresponding profile generated by the ad exchange. This map can then be used to select pollution content. Unlike re-marketing based pollution, the impact of behavioral pollution on altering user profiles towards more lucrative advertisers depends on the users' existing online interest profile. We empirically evaluate this impact across diverse user profiles in Section 5.

## 3.2 Hosting and Loading Pollution Content

The pollution content hosted by the fraudulent publisher should be loaded by the user's browser in a way that is not discernible by the user and ad exchange. While there are many ways to fabricate such camouflaged requests, such as `CSRF`, `XSS` and `Clickjacking` etc., in this paper we assume the pollution content is loaded using cross reference issued by hidden HTML `iframes`. These `iframes` are located outside the viewing area of the browser or layered underneath other content, and are used to reference and load pollution content. The loading of such content takes place in the background and is completely hidden from the user. Moreover, since approaches for frame-busting are not ubiquitously deployed, simple approaches can be used to hide the frame content from web crawlers.

**Embedding third-party websites.** Websites that want to prevent being embedded within an `iframe`, often as means to mitigate

[1]DoubleClick itself provides instructions on how their Tag Assistant detects re-marketing scripts. For more details, see https://support.google.com/tagassistant/answer/2954407?hl=en

clickjacking [27], employ techniques such as X-Frame-Options HTTP response header or Frame-Busting. We study the prevalence of these techniques by crawling the top 500 webpages belonging to each of the top 20 Alexa sub-categories. For each website, we tested whether it uses X-Frame-Options or one of the known Frame-Busting techniques [34].

Figure 2 shows the percentage of websites that use X-Frame-Options and Frame-Busting ordered by the aggregate percentage of the two methods. Only 5 sub-categories have more than 15% of the top websites that deploy embedding protection techniques. The vast majority of categories have less than 5% of their websites that employ such techniques, and the average across all is 4.6%. This shows that X-Frame-Options and frame busting are not ubiquitously deployed and the attack can leverage pollution content across a wide range of categories.

**Avoiding detection from web crawlers.** Hosting pollution content can have many adverse effects if detected by search engine crawlers. Upon detecting content embedded in hidden iframes, search engine crawlers can potentially flag the pollution content as malicious, blacklist the website or even ban the website from the search engine's index [11]. Nevertheless, it is possible to circumvent detection from web crawlers by generating pollution content dynamically in an obfuscated JavaScript code block, similar to how malicious websites that host drive-by download scripts evade scans from security checking web crawlers [19, 28]. Specifically, the fraudulent publisher can use obfuscated JavaScript code to show crawlers benign content rather than the pollution content available to real users.

## 3.3 Attack Victims

The primary victims of the attack are the advertisers that are being scammed to bid higher for users that are not really interested in their offering. The other victims are the website visitors, who have their profiles altered as a result of the attack. The immediate result is that the ads a user sees are irrelevant to her real interests, which might be offensive in some contexts. A different outcome can be in cases where the user's online interest profile is also used for personalizing other services. For example, Google may potentially use the same user profile to recommend movies on YouTube and even re-order search results based on user's online interests [9].

## 3.4 Attack Monetization – CPM and CPC

An important property of the attack is that it can be used to further boost the revenue generated by existing click and impression fraud mechanisms. This can be achieved if the bot master has control over the user's browser such that it can pollute user profiles to maximize the impact of the fraud. When deployed in isolation of existing fraud mechanisms, the attack is most effective for CPM-based ad campaigns. This is because CPM-based campaigns, which are the most common campaigns for display ads [25, 35], provide consistent cash flow to the publisher, regardless of whether visitors click on the potentially unrelated ads. In the rest of the paper we focus on CPM-based campaigns and assume that the attack is deployed as a standalone attack without deploying additional fraud methods.

# 4. ATTACK SETUP

In setting up the profile pollution attack we seek to address two main objectives. First, the attack setup should validate the complete end-to-end attack. Second, the attack setup should enable a detailed characterization of the effectiveness of the attack. Ideally, this can be achieved by compromising a legitimate publisher website to host pollution content, polluting the profile of users visiting the compromised publisher website, and monitoring the change in ad revenue generated by the publisher. However, doing so in a live setting raises several ethical concerns. Moreover, it limits our ability to provide a detailed characterization of the attack since it is not feasible to record the ads served to real users without the cooperation of the ad exchange.

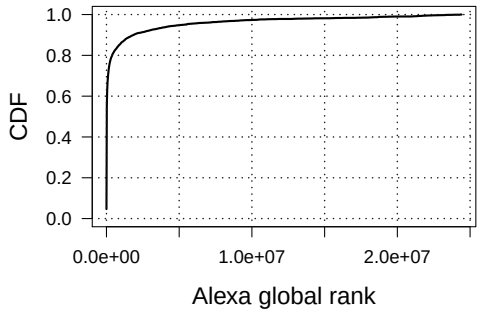



Figure 3: Cumulative distribution of the Alexa ranking of domains in the web traces.

To this end, we set up the attack as follows. Instead of driving live traffic, we emulate users browsing the websites with web traces. A few domains from the users' traces are selected as the fraudulent publishers. As we do not have control over these websites, the profile polluter is separated from the fraudulent publisher and is responsible for polluting the emulated user traffic immediately after loading the publisher's page to approximate a publisher that pollutes his own users. A distributed testbed of 200 nodes spread across the world (using PlanetLab) is used to generate web traffic to ensure location diversity. Since the traffic is emulated from browsers that we control, an ad crawler is used to record all the DoubleClick ads delivered to the emulated users. The recorded ads are analyzed and the revenue is estimated using publicly available CPM index values published by DoubleClick [26]. We also set up our own website as a fraudulent publisher to characterize the effectiveness of the attack.

## 4.1 User Web Traces & Profiles

Our attack setup replays *complete* web traces from real users to characterize and validate the attack. This is important because the ad revenue is not only impacted by the frequency with which users visit the publisher page but also depends on the user's online interest profile before and after pollution; the pollution impact depends on websites visited *prior* to pollution and the duration of the impact depends on websites visited *after* pollution.

We use web traces of real users from a Chrome extension installed by more than 700 users who have been using the extension for 2 years for research purpose. The functionality of the extension was modified to record all the webpage URLs visited by the user for a one week period (March 10th, 2014 - March 16th, 2014)[2]. In this time period, we collected a total of 224,855 page visits from 619 unique active users. Our dataset is diverse and consists of users using the extension across the world.

The web traces we use consist of users with diverse online interests and browsing behavior that are located across the world. Figure 3 shows that the websites in our dataset cover a large range of Alexa ranking, from extremely popular websites such as `google.com` and `facebook.com`, to websites that are ranked very low. Fig-

[2]IRB approval was granted and users were notified about the type of data collected and the intent of use for research purposes

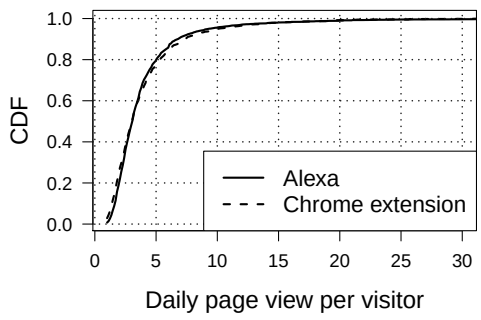



Figure 4: Cumulative distributions of the number of daily visits per user to the same domain in our web traces v.s. that of Alexa top domains.

ure 4 shows the average daily page view by a user for a given domain, exhibiting a similar wide range, from websites that have one visit per visitor to 10s of pages per visitor. To validate that visit patterns in our dataset are not skewed, the Figure also plots the distribution of daily page views per visitor for the top-100 websites in each of 200 Alexa categories. We observe that the two distributions almost completely overlap.

We create multiple copies of each user's profile to load under different pollution settings. The user web traces are replayed to generate profiles that are polluted by forwarding the request to the profile polluter after visiting the fraudulent publisher. We also create clean profiles by replaying user web traces bypassing the profile polluter. To eliminate the impact from time and location on distribution of ads, each user's profiles are generated by replaying her web trace at the same time from the same IP address. This provides a seamless approach to measure the extent to which the pollution impacts the type of ads targeted at the user. Thus, for every experiment presented in the following two sections, we record the ads targeted at the user with and without profile pollution.

## 4.2 Pollution Content

As described in Section 3.1, user profiles are polluted in order to mislead the ad exchange and advertisers from more lucrative verticals to target ads at users visiting the fraudulent publisher's webpage. We pollute each user by generating camouflaged visits to three websites from the top three most expensive display ad categories of *Health*, *Business* and *Education*. Beyond the top three ad categories, we pick two additional categories of *Sports* and *Shopping* to study the attack on less valuable ad categories.

**Polluting for Behavioral Targeting.** In order to find websites that alter the user's interest profile towards the above mentioned categories, we first filter websites from the corresponding Alexa category that contain the DoubleClick tracking script. For each website in this list, we use the Google Ad Preferences Dashboard [7] to build a map between the websites and categories that are consistent with DoubleClick. Table 1 provides the three websites selected for each category.

**Polluting for Re-Marketing Targeting.** Similar to the previous approach, we filter websites in the Alexa category that host the re-marketing script from DoubleClick. In addition to verifying that the category matches, we also verify that the re-marketing campaign is active. Table 1 lists the websites used for re-marketing pollution for each category.

## 4.3 Publisher Webpage

The complete attack is validated on two different type of publisher webpages.

**Live Websites.** We validate the attack on existing live publishers whose ad revenue is impacted by the dynamic content hosted by them as well as pre-existing preferences about type of ads that are allowed to be targeted. To this end, we select the top 19 most visited websites that host DoubleClick ads from the user web traces. Instead of compromising these websites to host pollution content, we set up the profile polluter as a separate entity. When emulating traffic traces, we forward the user to the profile polluter immediately after visiting any one of these 19 websites. We use results from these publishers primarily to estimate the revenue generated by the attack (Section 6).

**Controlled Publisher.** In order to form a baseline of the effectiveness of the attack, we set up our own publisher website and sign up with AdSense [2]. The publisher website has two display ad slots (top banner display ad and a side display ad) and uses the default settings provided by AdSense. Since AdSense requires the website to host some content before approving it, we upload static content that describes the different ad targeting mechanisms available to advertisers. Visiting the webpage with a blank profile results in DoubleClick profiling the user with interests belonging to the *Computers* category. Similar to the above setup, the profile polluter is separated from the controlled publisher.

## 4.4 Trace Emulator and Ad Crawler

A critical component of the attack setup is a distributed infrastructure to emulate web traffic by replaying the traces and recording all the ads delivered to the emulated user.

### 4.4.1 *Trace Emulator*

We develop a distributed infrastructure based on the PlanetLab testbed that is able to emulate real user web traffic. The trace emulator consists of a central control server and 264 worker nodes distributed across the world. The server maintains a list of tasks that are fetched by distributed workers periodically. Given a task containing the URL to visit and a unique user ID, the worker node instance loads one profile of the corresponding user, visits the assigned URL, records all the ads displayed on the webpage and the associated metadata, and sends this information back to the central server. The user's profile is updated accordingly after visiting the assigned URL.

### 4.4.2 *Ad Crawler*

Collecting measurements about display ads requires the ability to disassemble the elements of a webpage, identify ad elements and associate these with particular categories. Existing ad monitoring and blacklisting tools – AdBlock [1] and Ghostery [6] – work by matching URL patterns embedded in a webpage against a set of blacklist patterns, and cannot look deeper into the element and reason about it. The task is made even more difficult by complex DOM structures, deep nesting of elements, and dynamic JavaScript execution, that is found on a large fraction of pages on the Internet today. To address these challenges we extend the PhantomJS headless browser[3] to reliably extract the ad elements of a page, identify the actual landing pages for the ad elements, and associate the ads with specific semantic categories. The current implementation of the ad crawler is limited to ads delivered by DoubleClick. In the

[3]http://phantomjs.org

Table 1: The websites we use for polluting users' profiles in the five ad categories.

| Google Category | Alexa Category | Re-marketing Pollution Contents | Behavioral Pollution Contents |
|---|---|---|---|
| Health | Health | *eyemagic.net* | *intensemuscle.com*<br>*bimabazaar.com*<br>*allacqueredup.com* |
| Business | Business | *incorporate.com* | *bloomberg.com/news/insurance/*<br>*bloomberg.com/news/finance/*<br>*bloomberg.com/news/industries* |
| Educaton | Reference | *asuonline.asu.edu* | *universando.com*<br>*campusleader.com*<br>*graphs.net* |
| Shopping | Shopping | *teleflora.com* | *alterationsneeded.com*<br>*modernsalon.com*<br>*viloux.com* |
| Sports | Sports | *moenormangolf.com* | *bloguin.com*<br>*retospadel.com*<br>*golftechnic.com* |

following, we present an overview of the main modules of the ad crawler.

**DOM Parser/Preprocessor.** This module parses the DOM structure of the page and extracts specific attributes of display ads that reveal the landing page for the ad (the website that would be visited by clicking on the ad). This is complicated by the fact that display ads are often embedded in nested `iframe` tags spanning multiple levels[4]. In order to bypass the *same origin policy* enforced by modern web browsers we disable the web security mechanism of the PhanromJS headless browser. The DOM Parser reads the `<href>` (or `<flashvars>`) attributes for image (or flash) ads, and aggregates this information for further processing.

This module also logs DoubleClick elements (re-marketing scripts and cookies) on the page. Re-marketing scripts are detected by searching for the unique DoubleClick JavaScript code as described here [10]. DoubleClick cookies are detected by monitoring outgoing HTTP requests and comparing against the publicly available patterns provided by the Ghostery [6] tracker database.

**Ad Landing Page Extractor.** For each identified ad element, this infers the landing page by parsing the value of the attributes extracted by the DOM parser module and searching for specific patterns in the URL like `adurl=`, `redirect_url=`, etc. We manually generate these patterns for DoubleClick by inspecting the attribute value. In our experiments, we found that almost 80% of the ads have a landing page that is encoded by these patterns, while the remaining ads require actively following HTTP redirects. We do not follow these redirects; doing so could artificially inflate the click through rates of the ad campaigns, and bias the user profile inferred by DoubleClick towards these ad categories.

**Semantic Categories of the Ad URL.** The landing domains of the ads are categorized into one of the 13 top-level Alexa web categories. For each of the landing domain we collected, we queried `www.alexa.com` to find its corresponding category and assign the ad landing domain to one or more corresponding Alexa categories[5]. For those landing domains with no results returned from Alexa, we query `www.similarsites.com` as it provides a similar categorizer using the Alexa taxonomy for a much larger catalog of URLs[6]. Finally, if the query fails on both services, we manually categorize the URL.

**Verifying the Ad Crawler.** We verify the end-to-end implementation of the ad crawler by performing the following test. We set up our own re-marketing ad campaign on Google AdWorlds and drive fake traffic to the website using the traffic emulator. During the period that our campaign was active, we verified that the ad crawling framework is indeed able to capture our re-marketing ads across a wide variety of websites and the number of ad impressions closely matched the number reported by Google AdWords.

## 4.5 Ad Revenue Estimation

The final component of our setup is to estimate the ad revenue the fraudulent publisher generates. We estimate the revenue by using the publicly available report provided by Google that ranks the CPM cost of different ad verticals (categories) and associates with each vertical a relative cost index [26]. In the Google report, the index for the three most expensive categories of `Health`, `Business` and `Job & Education` were 257, 221 and 200 correspondingly. The least expensive category was `Law & Government` with an index of 46. We further manually mapped each of the 13 top-level Alexa categories to one of the ad verticals used in the published Google report. Our revenue estimation analysis always compares the revenue generated by the pollution attack with the baseline revenue computed by running the exact same experiment without pollution.

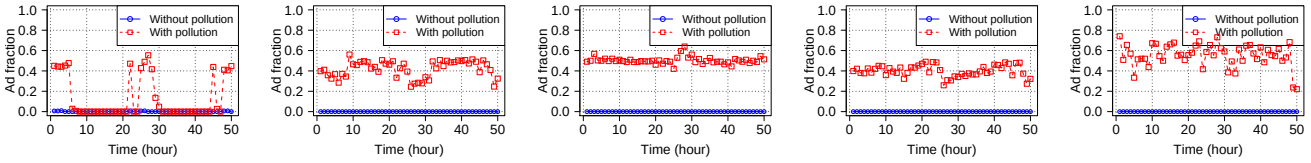
# 5. VALIDATION AND EFFECTIVENESS OF THE ATTACK

In this section we evaluate the extent to which profile pollution impacts the ads by deploying the attack on the controlled publisher webpage. The primary metric used for the evaluation is the relative change in ads from the desired ad category (behavioral pollution) or domain (re-marketing pollution) with and without pollution. For both user profile sets, the trace emulator first visits every website in a user's trace to ensure that all users have an online interest profile. We then take one set of user profiles and pollute all users only once. Subsequently, users from both sets visit the controlled publisher page once every hour for a duration of 50 hours. For each visit to the publisher webpage the ad crawler captures all the ads.

[4]In our experiments we observed up to six levels of nesting.

[5]We eliminate the top-level category of *Regional* as it provides information about the location of the URL

[6]We do not query this website directly as it is rate limited and requires purchasing an API access




(a) Sports (b) Reference (c) Business (d) Health (e) Shopping

Figure 5: Effectiveness of pollution attacks against re-marketing ad campaigns across different ad categories.

Table 2: Fraction of total ad impressions resulting from the re-marketing based pollution for varying number of re-marketing ad campaigns used for pollution.

| Number of Advertisers | Avg. | (Min, Max) |
|---|---|---|
| No pollution | 1.86% | (0.68%, 3.51%) |
| 1 | 55.55% | (22.06%, 74.07%) |
| 2 | 61.36% | (26.47%, 84.61%) |
| 3 | 80.11% | (39.71.0%, 97.40%) |
| 4 | 82.02% | (42.34%, 98.72%) |
| 5 | 87.87% | (53.29%, 100.00%) |

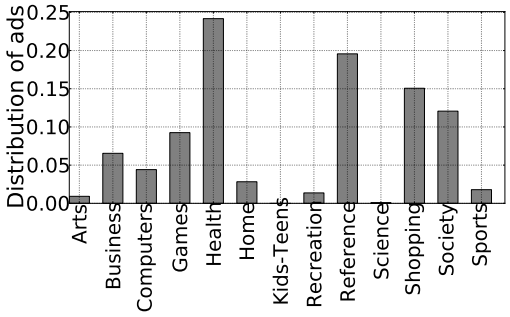



Figure 6: Distribution of ads across the 13 top-level Alexa categories.

## 5.1 Pollution Using Re-Marketing Campaigns

Figure 5 shows, for each pollution category, the fraction of ads received with and without profile pollution across all users. As expected, we observe that only the polluted users receive ads from the category used in the pollution. Surprisingly, we observe that across all categories, re-marketing ads aggressively target users, both in terms of time between the pollution and first ad shown, and number of ads: across *all* pollution categories users receive ads from the intended advertisers immediately in the very first visit to the publisher's webpage, and approximately 40–50% of all display ads are from these advertisers. We also verified the distribution of ads across users (not shown) and found that *all* the users received ads from the re-marketing campaigns used for pollution.

Next, we characterize the extent to which the polluter can increase the number of ad impressions served from the re-marketing advertisers by varying the number of advertisers used for pollution. Recall that the publisher website is set up with two display ads and hence requiring at least two re-marketing scripts to fill both slots. In practice, ad exchanges receive bids from other advertisers targeting the user and consequently not all ad slots may be filled by the re-marketing advertiser. Table 2 shows the fraction of re-marketing ads delivered to the users who are polluted by increasing number of re-marketing campaigns from one to five. We observe that with two re-marketing campaigns, the polluter can modify an average of 61% of ad impressions (min=26%, max=85%) and with three campaigns the average increases to 80% (min=40%, max=97%). Beyond three advertisers, the increase in fraction of ads delivered from the compromised campaigns is much lower. This indicates that the aggressive targeting performed by re-marketing campaigns enables the polluter to control a large fraction of ads displayed to the user.

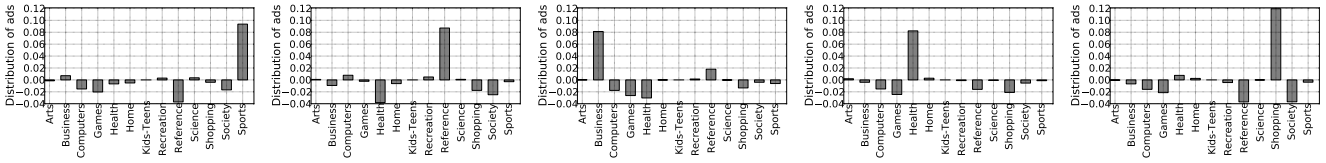
While the pollution is highly effective once it is triggered, advertisers may set up specific rules to trigger the campaign that can impact the publisher's ad revenue. First, as seen in Figure 5 advertisers can set up time based triggers. For example, the advertiser `moenormangolf.com` from the Sports category set up the campaign to only run during 8 hours of the day, causing a diurnal pattern in the targeted ads. Alternately, campaigns can be set up with frequency caps or they may be paused by the advertiser. Additionally, the advertiser may set up the campaign with a more complex control flow of user actions (*e.g.*, went to homepage, placed things in the cart, but never checked-out) and trigger the campaign only when a user completes all the steps. Thus, the primary challenge in effectively exploiting re-marketing campaigns is to select pollution content that accounts for such specific trigger rules.

## 5.2 Pollution Using Behavioral Targeting Campaigns

**Impact on Ad Categories.** We characterize how the pollution attack alters the distribution of ads across the different semantic categories. Intuitively, we expect that polluting a user's profile towards a specific category (*e.g.*, *Health*) should increase the number of ads delivered in that category. Correspondingly, since the total number of ad slots (and ad impressions) are the same, ads delivered from other categories should decrease. This relative change impacts the publisher's net revenue .

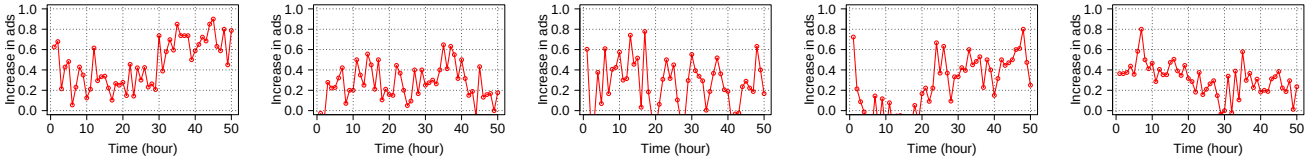
Figure 6 shows the distribution of ads across the 13 top-level Alexa categories for user profiles that were not polluted. We observe that the ad distribution spans multiple categories as users have diverse online interests. We use this baseline distribution and compute the relative change in the distribution of ad categories after pollution. Figure 7 shows the relative change in the ad categories across users which validates the effectiveness of our pollution attack – there is a clear increase in ads in the polluted category with a maximum increase of 12% for the *Shopping* category. Moreover, the pollution manages to increase the number of ads in categories that a user already received prior to pollution. For example, the fraction of ads in the *Health* category increases from 23% to 31%.

**Temporal Impact.** Finally, we study the temporal effect of the pollution. Figure 8 shows the relative increase in fraction of ads received from the category used for pollution. We observe that the effect of the pollution is immediate and leads to an increase in ads from the desired category. Moreover, the effect of the pollution persists over the entire time duration of the experiment. This indicates




(a) Sports (b) Reference (c) Business (d) Health (e) Shopping

Figure 7: Change in the distribution of ads.




(a) Sports (b) Reference (c) Business (d) Health (e) Shopping

Figure 8: Percentage increase in ads (pollution - no pollution) from the polluted category.

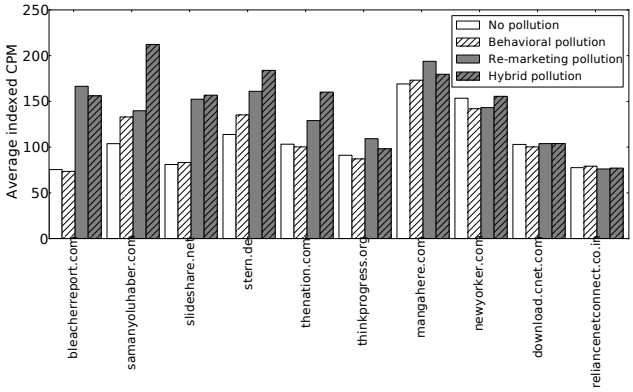



Figure 9: Average indexed CPM across the top 5 and bottom 5 selected websites before and after pollution.

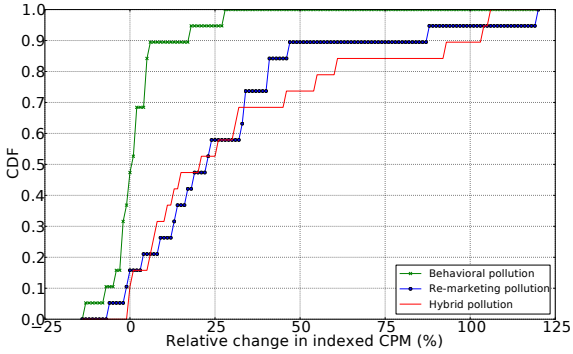



Figure 10: Distribution of the relative increase in the indexed CPM across the 19 selected websites.

that categories introduced artificially as an effect of the pollution have a lasting influence on the ads received by the user.

**Summary of behavioral pollution.** The results above indicate that our approach for behavioral pollution does indeed impact the ads targeted at the user while the effectiveness of the pollution depends on many factors like the user's existing profile, context of the publisher webpage as well as the category used for pollution. Unlike re-marketing pollution, the success of behavioral pollution is not dependent on a relatively small number of specific campaigns. Consequently, this makes it challenging for the fraudulent publisher to predict the exact landing domains and number of behavioral ads that will be served by the ad exchange.

Despite these sources of variability for re-marketing and behavioral pollution, it is still feasible for the fraudulent publisher to significantly increase its ad revenue.

# 6. REVENUE ESTIMATION FOR LIVE PUBLISHERS

In this section, we deploy the attack on live publisher websites and estimate the revenue generated by the attack for these publishers. Unlike the controlled publisher setting, there are a number of factors like the hosted content, popularity of the website, and ad preferences setup by the publisher that impact the ad revenue. While it is not feasible to explain the specific factors that impacts the publisher's revenue, we seek to empirically measure the overall impact of the pollution on the revenue of live publishers.

As described in Section 4.3, we select the top 19 most frequently visited websites from the web traces that host DoubleClick ads as the "fraudulent" publishers. When replaying the web traces, every visit to one of these 19 domains is followed by visiting the profile polluter. We emulate these traces four times in parallel for the following four pollution configurations - without pollution, behavioral pollution, re-marketing pollution, and hybrid (both) pollution using the pollution content shown in Table 1. The revenue is estimated using the CPM index [26] data reported by DoubleClick.

## 6.1 Aggregate CPM Index Change

Figure 10 shows the relative change in the CPM index for the three pollution configurations across the 19 websites. Overall, we find that behavioral pollution is not as effective as re-marketing based pollution; for almost 80% of the websites the change in the indexed CPM is not significant ($\pm$ 5%). On the other hand, re-marketing based pollution does significantly and consistently increase the relative indexed CPM; an increase of 4–120% for about 80% of the domains.

To better understand these distributions, Table 3 provides the traffic statistics along with the relative change of CPM index for the top five and bottom five performing domains ordered by the CPM

Table 3: Details of revenue experiments, showing the top 5 and bottom 5 websites we designated as fraudulent publishers ranked by relative change in indexed CPM using profile pollutions.

| Site | Alexa global rank | Avg page views per user per day | Num users | Avg page views per day | Change (% behavioral) | Change (% re-marketing) | Change (% hybrid) |
|---|---|---|---|---|---|---|---|
| *bleacherreport.com* | 231 | 3.96 | 133 | 527 | -2.60 | 120.64 | 106.81 |
| *samanyoluhaber.com* | 1,396 | 13.44 | 85 | 1142 | 28.11 | 34.67 | 104.52 |
| *slideshare.net* | 120 | 2.29 | 146 | 335 | 2.78 | 88.15 | 93.57 |
| *stern.de* | 1,691 | 2.58 | 60 | 155 | 18.75 | 41.43 | 61.54 |
| *thenation.com* | 13,835 | 1.50 | 88 | 132 | -2.88 | 24.99 | 55.15 |
| *thinkprogress.org* | 3,960 | 1.37 | 91 | 125 | -4.37 | 19.90 | 7.70 |
| *mangahere.com* | 1,903 | 72.71 | 52 | 3781 | 2.43 | 14.63 | 6.25 |
| *newyorker.com* | 2,432 | 1.93 | 159 | 307 | -7.49 | -6.67 | 1.37 |
| *download.cnet.com* | 104 | 4.48 | 69 | 309 | -2.62 | 0.86 | 0.95 |
| *reliancenetconnect.co.in* | 1,694 | 1.79 | 102 | 183 | 2.07 | -1.85 | -0.61 |

index with hybrid pollution. Figure 9 shows the average indexed CPM for the same 10 websites. We make a number of observations from this data:

**Website Ranking and Traffic Patterns.** Across the five best and worst performing websites we do not observe any correlation between the website ranking or traffic patterns with the revenue generated by either one of the three pollution configurations. This indicates that our attack is able to deceive the ad exchange in targeting high value ads even on websites that are ranked much lower or have highly varying traffic patterns.

**Varying Performance of Behavioral Pollution.** We observe that behavioral pollution does not consistently increase the ad revenue for the fraudulent publisher. Among the top five websites listed in Table 3, *bleacherreport.com*, *slideshare.net* and *thenation.com* yield a negative or very low increase in the average CPM index. Looking into the logs, we find that the behavioral pollution of the emulated traffic to these websites was ineffective. For example, 83% and 85% of the ads targeted on *bleacherreport.com* were from a single advertiser, *ford.com*, before and after behavioral pollution, respectively. Similarly, 100% and 93% of the ads on *slideshare.net* were from *academy.com* before and after behavioral pollution, respectively. On the other hand, re-marketing and hybrid pollution for these domains was effective and led to a significant increase in ad revenue. This potentially indicates that these websites have pre-sold their ad inventory and consequently behavioral pollution was ineffective. However, re-marketing based pollution manages to override this pre-sold ad inventory, potentially because of the higher CPM and CPC costs associated with these ads.

**Low Yield Re-marketing Pollution.** As discussed in Section 5, re-marketing based pollution leads to aggressive targeting of users independent of their online profile. However, we observe that for *newyorker.com*, *download.cnet.com*, and *reliancenetconnect.co.in* all three pollution configurations are ineffective. None of the three domains received ads from the advertisers used for re-marketing pollution even when users visiting other domains were targeted with the re-marketing ads. Moreover, the behavioral pollution was also ineffective for these domains. For example, on *reliancenetconnect.co.in*, between 65%-73% of the ads targeted at users before and after pollution (all three pollution types) were automobile related ads from domains like *mazdausa.com*, *avis.com*, *budget.com* and *driveamazda.com*. This potentially indicates a scenario where the publisher website is explicitly configured to only receive automobile related ads making the different pollution mechanisms ineffective.

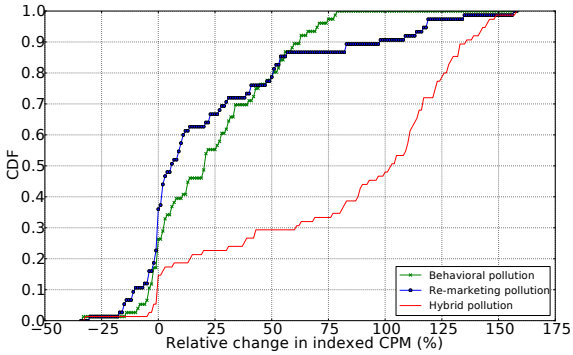



Figure 11: Distribution of the relative change in the average indexed CPM across the visitors of *samanyoluhaber.com* as a result of the pollution attack.

## 6.2 Revenue Contribution Per User

Next, we seek to understand how individual users contribute to the revenue of a fraudulent publisher. To this end, Figure 11 shows the cumulative distribution of the relative change of the indexed CPM for each of the 85 visitors of *samanyoluhaber.com*. The overall distribution is biased towards a positive increase in the indexed CPM influencing the average value of the improvement in Table 3 for the *samanyoluhaber.com*. We observe a large variation in the relative CPM index for all three pollution configurations; the relative indexed CPM can reduce by as much as ∼30% and increase by up to 79%, 157% and 158% for behavioral, re-marketing and hybrid pollution, respectively. This wide range is expected as users' online interests and browsing patterns vary, and this impacts the extent to which higher-paying ads are targeted after pollution.

We observe that across the three pollution configurations the relative indexed CPM decreases for 17.1%, 22.7% and 5.3% of users for behavioral, re-marketing and hybrid pollution. Looking into the traces of these specific users, we observed that these users visit the publisher webpage infrequently with a median of only 2 visits to the publisher websites.

The profile pollution attack significantly increases the average indexed CPM ( $>$ 50%) for 22.4%, 22.6% and 70.6% of users for behavioral, re-marketing and hybrid pollution, respectively. For these users we observe the opposite trend; these users visit *samanyoluhaber.com* frequently, with a median of 9.5 visit across all users. This frequent visit pattern is ideal for the attack and enables the fraudulent publisher to repeatedly pollute the user's profile for each successive visit.

## 6.3 Summary of impact on ad revenue

Overall, our results indicate that the preferences set by the publisher when signing up with the ad exchange have a direct impact on the revenue generated by the profile pollution attack. Despite these preference, the pollution attack is *lucrative* and can indeed increase the publishers' revenue with an average of 2.34%, 29.62% and 33.89% for the three pollution configurations. Furthermore, this revenue increase is *robust* to diverse online profiles of users, ranking of the publisher webpage, and varying visit patterns to the publisher webpage.

# 7. COUNTERMEASURES

In this section we discuss countermeasures and best practices that different entities in the ad ecosystem can adopt in order to mitigate or at least minimize the attack surface.

## 7.1 Publishers

Commonly, websites are not supposed to be framed within another website as part of an `iframe` [12]. Therefore, using `X-Frame-Option` or deploying a "frame-busting" method can make it more difficult for the polluter to abuse innocent websites for the purpose of pollution fraud (other methods, such as pop-unders can still be used, but are easier to detect).

## 7.2 Advertisers

Advertisers should protect their ad campaigns against pollution attacks by targeting audiences that have very specific interests. This effectively raises the bar for the polluter to find relevant pollution content impacting a large number of users. For example, finding the appropriate pollution content for the category *Jobs & Education→Education→Distance Learning* may be more difficult to compared to finding pollution content for *Education*. Similarly, a re-marketing campaign that targets users with a specific flow in the website, *e.g.*, users who logged in, placed an item in a cart but did not check out, is more difficult to compromise compared to targeting all users who visited the webpage of the advertiser. We note that the downside of such fine-grained audience targeting is that it may reduce the size of the target audience.

## 7.3 Ad Exchange and Ad Networks

Recent work, like ViceROI [22], aims to detect click spam by comparing the revenue per user for a fraudulent publisher with a baseline set of ethical publishers. While this approach is limited to catching click spam, ad networks and ad exchances should deploy similar approaches to detect impression fraud caused by anomalous revenue changes in the fraudulent publisher's ad revenue. Eventhough the attacker has control over his ad revenue through configuring the attack settings (e.g. pollution content, ad preference, and amount of polluted users, etc.), the deployment of systems like ViceROI could reduce the ad revenue generated from profile pollution.

Ad exchanges like DoubleClick do not check for the domain in which the re-marketing script is being executed. Consequently, it is sufficient for the polluter to simply copy the JavaScript provided by the ad exchange. To prevent this, the re-marketing script provided by the ad exchange should be bound to the designated domain, and at runtime the script should verify that it indeed runs within the intended domain.

A few ad exchanges and ad networks provide users the ability to inspect and modify the inferred online interest profile or opt out of personalized ads [4, 5, 7, 13, 14]. However, users have no visibility into how these profiles are generated or used to serve targeted ads [29]. Ad exchanges and ad networks should provide users easy mechanisms to flag suspicious ads they see that are not aligned with their real interests. Additionally, ad exchanges should also encourage users to manually adjust their online interests, and explicitly avoid being targeted in some categories. For example, a user might want to disallow all *Health* related ads. In such a case, a polluter attempting to influence the user's profile with the *Health* category would lead to no ads from this category to be targeted at the user.

A key contributor to the success of the attack is that pollution content immediately impacts the user profile, thus the polluter can almost immediately benefit from the attack. The ad networks can increase the duration between page visits and the impact on the user's profile, thus mitigating the impact of the attack by profiling users interests across a large set of websites visted by the user. However, this delay might be in contrast to the ad networks' desire for accurate and timely inference of user interests, especially for re-marketing campaigns.

# 8. RELATED WORK

*Online Advertising Economy and Tracking.* The economics of online advertising is discussed in detail in [23], which considers the usage of targeting users based on interests as a key difference between traditional and online advertising. More recently, Gill *et al.* [25] proposed a simple model for capturing the effect of user profile (or "intent") on the revenue obtained by the ad network and the publishers. Using this model the authors stressed the significance of the user profile in the ecosystem by showing that incorporating mechanisms that block tracking, thereby essentially eliminating targeted advertising, can decrease the overall revenue of ad networks by 75%.

In order to build an accurate user profile, ad networks need to track users as they browse the web. Several recent papers measure the extent to which users are being tracked and targeted by ad networks [29,30,33]. Rosner *et al.* [33] showed that online tracking of users is ubiquitous and covers a large fraction of a user's browsing behavior. Liu *et al.* [29] focused on Google's DoubleClick network and showed that interest-based targeting is prominent and spans multiple ad categories, with up to 65% of the ad categories received by a user are targeted based on the user's inferred profile.

In this work we leverage the strong relation between the user interest profile and the economics of online advertising to propose a method for polluting user interest profile for increasing the publisher's revenue.

*Pollution Attacks.* Pollution attacks against users have been shown to be useful for a variety of attacks, including influencing product recommendations on Amazon and video recommendations on YouTube [39]. While these attacks were shown to be useful in influencing the way users interact with the polluted system, it is not clear whether the polluter can actually gain monetary benefits from the attack. Our work is the first application of a pollution attack that leads to clear monetary value to the attacker.

*Fraud in Online Advertising.* Fraud in online advertising and countermeasures against these fraud mechanisms have been the focus of a long line of research efforts [8,16,17,20,21,31,35–38]. The most common fraudulent activities include those where fraudulent publishers leverage click-spam networks or pay-per-view networks to increase the traffic to their sites, and thus increase their ad revenue. Click-spam networks cause fraudulent clicks on ads in order to increase the income of the publisher or sometimes deplete the budget of the advertiser. The most recently study by [21], where the authors conducted a controlled experiment show that click-spam attacks account for 10–25% of the clicks, highlighting the prominence of such attacks. In a recent study, the authors used these

results and presented a system [22] that ad networks can use for catching click-spam in search ad networks.

Different from click-spam networks are pay-per-view networks that artificially increase the number of ad impressions of fraudulent publishers by framing the publisher's website within other websites in a camouflaged fashion. Fraudulent activities using pay-per-view networks typically result in impressions that are registered on the camouflaged pages without "genuine user interest" *i.e.*, invalid traffic generation. A recent study [35] have shown a pay-per-view network generates hundreds of millions of fraudulent impressions per day.

Existing online advertising frauds focus solely on increasing the volume of ad clicks or impressions and have largely ignored the impact of user profiles. Our attack complements these existing fraud mechanisms by enabling the publisher to further boost the revenue obtained by participating in either of the networks. Compared to existing fraudulent activities which are suspect to traffic analysis [35, 40], our attack is more resilient against current fraud detection methods.

## 9. CONCLUSION

This paper presents a new pollution attack on online targeted advertising that exploits the role played by the online interest profile of users in the ad selection process. The attack leverages novel mechanisms to pollute the profile of users visiting a publisher page in a way that deceives the ad exchange and advertisers to target more lucrative ads, thereby increasing the publisher's revenue. The proposed attack is validated and characterized for the two most commonly used ad targeting mechanisms (re-marketing and behavioral targeting) by emulating a real world deployment. The study shows that the profile pollution based attack is robust against diverse browsing patterns and online interests of users, and effective in drawing the intended higher-paying ads resulting in a significant increase in ad revenue.

## 10. ACKNOWLEDGMENTS

This material is based upon work supported in part by the National Science Foundation under Grants No. CNS-1017265, CNS-0831300, and CNS-1149051, by the Office of Naval Research under Grant No. N000140911042, by the Department of Homeland Security under contract No. N66001-12-C-0133, and by the United States Air Force under Contract No. FA8650-10-C-7025. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, the Office of Naval Research, the Department of Homeland Security, or the United States Air Force.

## 11. REFERENCES

[1] Adblock Plus. http://adblockplus.org/.

[2] AdSense. www.google.com/adsense/.

[3] AdSense revenue share. https://support.google.com/adsense/answer/180195?hl=en.

[4] Amazon.com: Advertising Preferences. http://www.amazon.com/gp/dra/info.

[5] Facebook Ads. https://www.facebook.com/settings?tab=ads&view.

[6] Ghostery. http://www.ghostery.com/.

[7] Google Ad Preferences Manager. https://www.google.com/ads/preferences.

[8] Google AdSense - Working better together: Protecting against invalid activity. http://adsense.blogspot.com/2012/12/working-better-together-protecting.html.

[9] Google Privacy Policy. http://www.google.com/policies/privacy/.

[10] Google Tag Assistant. https://support.google.com/tagassistant/answer/2954407?hl=en.

[11] Google Webmaster Guidelines - Quality. https://support.google.com/webmasters/answer/35769.

[12] Google Webmaster Tools - Frames. https://support.google.com/webmasters/answer/34445?hl=en.

[13] Microsoft personalized ad preferences. http://choice.microsoft.com/en-us/opt-out.

[14] Yahoo Ad Interest Manager. http://info.yahoo.com/privacy/us/yahoo/opt_out/targeting/details.html.

[15] The Value of Behavioral Targeting. http://www.networkadvertising.org/pdfs/Beales_NAI_Study.pdf, 2009.

[16] S. A. Alrwais, A. Gerber, C. W. Dunn, O. Spatscheck, M. Gupta, and E. Osterweil. Dissecting ghost clicks: Ad fraud via misdirected human clicks. In *Proceedings of the 28th Annual Computer Security Applications Conference*, ACSAC '12, pages 21–30, New York, NY, USA, 2012. ACM.

[17] V. Anupam, A. Mayer, K. Nissim, B. Pinkas, and M. K. Reiter. On the security of pay-per-click and other web advertising schemes. In *Proceedings of the Eighth International Conference on World Wide Web*, WWW '99, pages 1091–1100, New York, NY, USA, 1999. Elsevier North-Holland, Inc.

[18] A. Barth, C. Jackson, and J. C. Mitchell. Robust defenses for cross-site request forgery. In *Proceedings of the 15th ACM Conference on Computer and Communications Security*, CCS '08, pages 75–88, New York, NY, USA, 2008. ACM.

[19] M. Cova, C. Kruegel, and G. Vigna. Detection and analysis of drive-by-download attacks and malicious javascript code. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 281–290, New York, NY, USA, 2010. ACM.

[20] N. Daswani, C. Mysen, V. Rao, S. Weis, K. Gharachorloo, and S. Ghosemajumder. Online advertising fraud. *Crimeware: understanding new attacks and defenses*, 2008.

[21] V. Dave, S. Guha, and Y. Zhang. Measuring and fingerprinting click-spam in ad networks. *SIGCOMM Comput. Commun. Rev.*, 42(4):175–186, Aug. 2012.

[22] V. Dave, S. Guha, and Y. Zhang. Viceroi: Catching click-spam in search ad networks. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer and Communications Security*, CCS '13, pages 765–776, New York, NY, USA, 2013. ACM.

[23] D. S. Evans. The Economics of the Online Advertising Industry. *Review of Network Economics*, 7(3):359–391, 2008.

[24] A. Farahat and M. C. Bailey. How effective is targeted advertising? In *Proceedings of the 21st International Conference on World Wide Web*, WWW '12, pages 111–120, New York, NY, USA, 2012. ACM.

[25] P. Gill, V. Erramilli, A. Chaintreau, B. Krishnamurthy, K. Papagiannaki, and P. Rodriguez. Follow the money: Understanding economics of online aggregation and advertising. In *Proceedings of the 2013 Conference on Internet Measurement Conference*, IMC '13, pages 141–148, New York, NY, USA, 2013. ACM.

[26] What's Trending in Display for Publishers? http://www.google.com/think/research-studies/whats-trending-in-display-for-publishers.html.

[27] L.-S. Huang, A. Moshchuk, H. J. Wang, S. Schechter, and C. Jackson. Clickjacking: Attacks and defenses. In *Proceedings of the 21st USENIX Conference on Security Symposium*, Security'12, pages 22–22, Berkeley, CA, USA, 2012. USENIX Association.

[28] C. Kolbitsch, B. Livshits, B. Zorn, and C. Seifert. Rozzle: De-cloaking internet malware. In *Proceedings of the 2012 IEEE Symposium on Security and Privacy*, SP '12, pages 443–457, Washington, DC, USA, 2012. IEEE Computer Society.

[29] B. Liu, A. Sheth, U. Weinsberg, J. Chandrashekar, and R. Govindan. Adreveal: Improving transparency into online targeted advertising. In *Proceedings of the Twelfth ACM Workshop on Hot Topics in Networks*, HotNets-XII, pages 12:1–12:7, New York, NY, USA, 2013. ACM.

[30] J. R. Mayer and J. C. Mitchell. Third-party web tracking: Policy and technology. In *Proceedings of the 2012 IEEE Symposium on Security and Privacy*, SP '12, pages 413–427, Washington, DC, USA, 2012. IEEE Computer Society.

[31] A. Metwally, D. Agrawal, and A. E. Abbadi. Using association rules for fraud detection in web advertising networks. In *Proceedings of the 31st International Conference on Very Large Data Bases*, VLDB '05, pages 169–180. VLDB Endowment, 2005.

[32] F. Nentwich, N. Jovanovic, E. Kirda, C. Kruegel, and G. Vigna. Cross-site scripting prevention with dynamic data tainting and static analysis. In *In Proceeding of the Network and Distributed System Security Symposium*, 2007.

[33] F. Roesner, T. Kohno, and D. Wetherall. Detecting and defending against third-party tracking on the web. In *Proceedings of the 9th USENIX Conference on Networked Systems Design and Implementation*, NSDI'12, pages 12–12, Berkeley, CA, USA, 2012. USENIX Association.

[34] G. Rydstedt, E. Bursztein, D. Boneh, and C. Jackson. Busting frame busting: a study of clickjacking vulnerabilities at popular sites. In *In IEEE Oakland Web 2.0 Security and Privacy Workshop*, page 6, 2010.

[35] K. Springborn and P. Barford. Impression fraud in online advertising via pay-per-view networks. In *Proceedings of the 22Nd USENIX Conference on Security*, SEC'13, pages 211–226, Berkeley, CA, USA, 2013. USENIX Association.

[36] O. Stitelman, C. Perlich, B. Dalessandro, R. Hook, T. Raeder, and F. Provost. Using co-visitation networks for detecting large scale online display advertising exchange fraud. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '13, pages 1240–1248, New York, NY, USA, 2013. ACM.

[37] B. Stone-Gross, R. Stevens, A. Zarras, R. Kemmerer, C. Kruegel, and G. Vigna. Understanding fraudulent activities in online ad exchanges. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, IMC '11, pages 279–294, New York, NY, USA, 2011. ACM.

[38] G. Wang, C. Wilson, X. Zhao, Y. Zhu, M. Mohanlal, H. Zheng, and B. Y. Zhao. Serf and turf: crowdturfing for fun and profit. In *Proceedings of the 21st international conference on World Wide Web*, WWW '12, pages 679–688, New York, NY, USA, 2012. ACM.

[39] X. Xing, W. Meng, D. Doozan, A. C. Snoeren, N. Feamster, and W. Lee. Take this personally: Pollution attacks on personalized services. In *Proceedings of the 22Nd USENIX Conference on Security*, SEC'13, pages 671–686, Berkeley, CA, USA, 2013. USENIX Association.

[40] L. Zhang and Y. Guan. Detecting click fraud in pay-per-click streams of online advertising networks. In *Proceedings of the 2008 The 28th International Conference on Distributed Computing Systems*, ICDCS '08, pages 77–84, Washington, DC, USA, 2008. IEEE Computer Society.