Betrayed by Your Ads!

Claude Castelluccia, Mohamed Ali Kaafar, Tran Minh-Dung

Betrayed by Your Ads!
Reconstructing User Profiles From Targeted Ads

Claude Castelluccia, Mohamed Ali Kaafar, and Minh-Dung Tran

Inria, France
{claude.castelluccia,mohamed-ali.kaafar,minh-dung.tran}@inria.fr

Abstract. In targeted (or behavioral) advertising, users' behaviors are tracked over time in order to customize served ads to their interests. This creates serious privacy concerns since for the purpose of profiling, private information is collected and centralized by a limited number of companies. Despite claims that this information is secure, there is a potential for this information to be leaked through the customized services these companies are offering. In this paper, we show that targeted ads expose users' private data not only to ad providers but also to any entity that has access to users' ads. We propose a methodology to filter targeted ads and infer users' interests from them. We show that an adversary that has access to only a small number of websites containing Google ads can infer users' interests with an accuracy of more than 79% (Precision) and reconstruct as much as 58% of a Google Ads profile in general (Recall). This paper is, to our knowledge, the first work that identifies and quantifies information leakage through ads served in targeted advertising.

Keywords: Targeted Advertising, Privacy, Information leakage

1 Introduction

Context: Internet users are being increasingly tracked and profiled. Companies utilize profiling to provide customized, i.e. personalized services to their customers, and hence increase revenues. In particular, behavioral advertising takes advantage from profiles of users' interests, characteristics (such as gender, age and ethnicity) and purchasing activities. For example, advertising or publishing companies use behavioral targeting to display advertisements that closely reflect users' interests (e.g. 'sports enthusiasts'). Typically, these interests are inferred from users' web browsing activities, which in turn allows building of users' profiles.

It can be argued that customization resulting from profiling is also beneficial to users who receive useful information and relevant online ads in line with their interests. However, behavioral targeting is often perceived as a threat to privacy mainly because it heavily relies on users' personal information, collected by only a few companies. In this paper, we show that behavioral advertising poses an additional privacy threat because targeted ads expose users' private data to any entity that has access to a small portion of these ads. More specifically, we show

that an adversary who has access to a user's targeted ads can retrieve a large part of his interest profile. This constitutes a privacy breach because, as illustrated in Section 2, interest profiles often contain private and sensitive information.

Motivation: This work was largely motivated by the Cory Doctorow's "Scroogled" short story that starts as follows [12]:

Greg landed at San Francisco International Airport at 8 p.m... The officer stared at his screen, tapping...

-"Tell me about your hobbies. Are you into model rocketry?"

-"What?"

-"Model rocketry."

-"No," Greg said, "No, I'm not."

-"You see, I ask because I see a heavy spike in ads for rocketry supplies showing up alongside your search results and Google mail."

-"You're looking at my searches and e-mail?"

-"Sir, calm down, please. No, I am not looking at your searches,... That would be unconstitutional. We see only the ads that show up when you read your mail and do your searching. I have a brochure explaining it ..."

The main goal of this paper is to study whether such scenario would be possible today, and if one can infer a user's interests from his targeted ads. More specifically, we aim at quantifying how much of a user's interest profile is exposed by his targeted ads. However, as opposed to the above story, we do not consider ads that show up when a user reads his email or uses a search engine. These ads are often contextual, i.e. targeted to email contents or search queries. Instead, we consider targeted ads that are served on websites when a user is browsing the web.

Contributions of this paper: We describe an attack that allows any entity that has access to users' targeted ads to infer these users' interests recovering a significant part of their interest profiles. More specifically, our experiments with the Google Display Network[4] demonstrate that by analyzing a small number of targeted ads, an adversary can correctly infer users' Google interest categories with a high probability of 79% and retrieve as much as 58% of Google Ads profiles.

The attack described in this paper is practical and easy to perform, since it only requires the adversary to eavesdrop on a network for a short period of time and collect a limited number of served ads.

The crux of the problem is that even if some websites use secure connections such as SSL (Secure Socket Layer), ads are almost always served in clear. For example, Google currently does not provide any option to serve ads with SSL[1] [2]. We acknowledge that in some scenarios the adversary can recover a user's profile directly from the websites he visits, i.e. without considering targeted ads. However, we show in this paper that targeted ads can often improve the accuracy of recovered profiles and reduce the recovery time. Furthermore, in

[1] We verified this feature by browsing through several https websites (e.g. https://www.nytimes.com/).

some circumstances, the victim has different browsing behaviors according to his environment. For example, a user at work mostly visits websites related to his professional activity, while he visits websites related to his personal interests at home. We show in this paper that an adversary, such as an employer, that can eavesdrop on the victim's computer or network while at work can infer information about his "private" and personal interest profile. In other words, targeted ads constitute a covert channel that can leak private information.

Although there are various targeted advertising networks today, this work focuses on Google advertising system, which is "the most prevalent tracker" according to a survey of The *Wall Street Journal* [17]. However, our methodology is general enough to be extended to other ad networks. The problem of generality will be discussed in Section 3.1.

Structure of the paper: Section 2 describes the Google targeted advertising system. In section 3, we present our approach to filter targeted ads and describe how we infer Google Ads profiles from them. We then present in Section 4 the performance of our method through some experiments in the Google Display Network. In section 5, we discuss the related work. Section 6 presents possible countermeasures and discusses some relevant problems. Section 7 concludes the paper.

2 Targeted Advertising: The case of Google

Google Display Network is a network of websites (also called publishers) that serve Google ads. Google receives ads from advertisers and then selects the appropriate publishers using various criteria such as relevant content, bid price and revenue.

In the Google targeted advertising model, Google Display Network sites are also used to track users as they browse the Internet. Each time a user visits a website that contains Google ads, i.e. a website that belongs to the Google Display Network, he sends his *DoubleClick*[2] cookie to Google, along with information about the visited website. As a result, Google collects all the sites within the Google Display Network that have been visited by a user, and builds an interest profile from them. A Google profile is defined as a set of categories and sub-categories (see figure 1). For example, if a user visits a football site several times, Google may assign him the category *Sport*, or more specifically the subcategory *Sport* $\rightarrow$ *Football*. In addition, a Google profile may include location information and some demographic data such as the gender and age of the user. These profiles are then used to target ads to users.

A user can access and modify his Google Ads Preferences by accessing the webpage `http://www.google.com/ads/preferences` [8]. Furthermore, a user can choose to opt out of the Google targeted advertising if he no longer wants to receive targeted ads. Figure 1 displays an example of a Google user profile

[2] In order to keep track of users visiting the Google Display Network, Google uses the DoubleClick cookie issued from the doubleclick.net domain which belongs to Google

Fig. 1. An Example of a Google Ads Preferences Page.

that contains potentially private and sensitive information. For example, the "Job listing" category indicates that the user is probably looking for a job. A user might probably want to keep this information secret, in particular from his current employer. Furthermore, the category "Dating & Personals" indicates that the user is currently actively looking for a relationship, and the subcategories "Baby names" and "Adoption" that he has been recently visiting web sites related to baby adoption.

In its privacy policy, Google states that "We take appropriate security measures to protect against unauthorized access to or unauthorized alteration, disclosure or destruction of data.", and "We restrict access to personal information to Google employees, contractors and agents who need to know that information in order to process it on our behalf" [5]. Nevertheless, in this paper we show that a portion of personal users' profiles could be leaked through targeted ads. Even if Google does not consider users' interests as "personal information", this data which is related to users online activities, can be very private from a user's perspective.

3 Inferring users' profiles from targeted Ads

As targeted ads are personalized to each user based on his profile, they can reveal a lot of information about users' interests. This section describes how an adversary who has access to an user's ads can derive part of his interests from them.

As shown in Figure 2, our approach is composed of two main phases. The first phase collects all ads served to a target user and filters them to only retain targeted ones. In the second phase, targeted ads are classified into categories

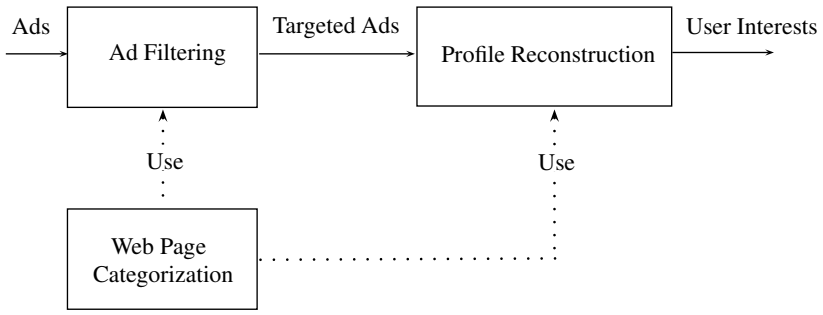

Fig. 2. Filtering targeted ads and inferring user interests.

and a profile is re-constructed. These two phases are detailed in the rest of this section. However, we first start by presenting two building blocks used by these both phases. The first one is an algorithm that is used to categorize webpages. The second one is a set of algorithms to compare categories.

3.1 Building Blocks

Web Page Categorization The web page categorization tool is an algorithm that derives interest categories from a web page. This algorithm relies on the tool used by Google to generate Google Ads Preferences pages. As described in the previous section, each time a user visits a web page, Google derives some interests in the form of categories and updates the user's Google Ads Preferences page accordingly.

We use this tool to build our page categorization algorithm. Given a webpage W, our algorithm operates as follows:

1. W is visited and and the resulting DoubleClick cookie is saved.
2. A request is made to the Google Ads Preferences page with the previously saved cookie. Note that Google does not always update the Google Ads Preferences page upon a single web page visit. Usually, a webpage has to be visited multiple times to update the Google Ads Preferences page. Furthermore, we noticed that users' Ads preferences are updated after a period of time ranging between 1 and 2 minutes. Therefore, this step is repeated 5 times (heuristic value) every 2 minutes.
3. The Google Ads Preferences page is parsed and the corresponding categories are retrieved.

To evaluate the performance of our approach, we scraped 5000 ads from Google search page and 2000 sites from Google Display Network, classified them by the tool, and reviewed the results. We detected that almost all of these pages can be categorized by Google (more than 90%). We also manually reviewed the categorization results and observed that, although there are some irrelevant

Ad Url	Categories
http://www.elasticsteel.net ID=156	Beauty & Fitness → Fitness
http://www.livecarhire.com	Travel → Car Rental & Taxi Services
http://www.terracebeachresort.ca	Travel → Hotels & Accommodations Travel → Tourist Destinations → Beaches & Islands
http://www.sanibelbayfronthouse.com	Arts & Entertainment → Entertainment Industry → Film & TV Industry → Film & TV Production Real Estate → Timeshares & Vacation Properties Travel → Tourist Destinations → Zoos-Aquariums-Preserves
http://www.siestakeyaccommodation.com	Real Estate → Timeshares & Vacation Properties Travel

Table 1. Ad page categorization example

categories, the categorization generally reflects the content of each page. Table 1 presents several examples of ad page categorization.

It should be noted that relying on Google does not reduce the generality of our method. There exist many efficient techniques to categorize the content of web pages. For example [16] uses cosine similarity. This method is very efficient since it relies on social/crowd data (folksonomy) which is continuously updated, and is appropriate for fine-grained categorization. We implemented this method and compared its performance with the Google-based categorization approach we described above. The obtained results were quite similar, with more than 60% of the categories overlapping. We therefore believe that our work can be extended to other ad networks, such as Yahoo! or Microsoft, either by applying their own categorization, or by simply using an existing webpages categorization technique. Note that Yahoo! and Microsoft also build users' behavior-based interest profiles and similarly to Google personalize ads to users according to their interests [9] [10].

Category Comparison Methods Many of the filtering and evaluation algorithms presented in this paper need to compare categories. We use three methods for this purpose: "Same category", "Same parent" and "Same root":

1. *Same category:* Two categories are considered equivalent in the "Same category" method if they match exactly.
2. *Same parent:* Two categories are considered equivalent in the "Same parent" method if they have the same parent category. For example, the two categories "Arts & Entertainment → Entertainment Industry → Film & TV Industry → Film & TV Awards" and "Arts & Entertainment → Entertainment Industry → Film & TV Industry → Film & TV Production" have

the same parent category "Film & TV Industry", so they are considered equivalent to each other in the "Same parent" method.

3. *Same root:* Two categories with same root category are considered equivalent in the "Same root" method. For example, the two categories "Arts & Entertainment → Entertainment Industry → Recording Industry → Music Awards" and "Arts & Entertainment → Movies → Action & Adventure Films → Superhero Films" have the same root category "Arts & Entertainment" and therefore are equivalent to each other in the "Same root" method. Obviously, if two categories are equivalent in the "Same parent" method, they are also equivalent in the "Same root" method.

3.2 Extracting Targeted Ads

Ads provided by Google are either location-based, content-based (we call hereafter contextual, i.e. related to the visited page's content), generic, or profile-based (we call hereafter targeted, i.e. customized to users' profiles). In this paper, we only consider targeted ads. We therefore need to filter out location-based, content-based and generic ads (see figure 3).

Fig. 3. Filtering targeted ads

We conducted all experiments with users from the same location. As a result, the location-based filter is not used (and therefore not presented here). Furthermore, we consider that an ad is contextual if it shares at least one category with its displaying page (the page on which the ad is delivered). To filter out contextual ads, we therefore categorize, using the categorization technique described in Section 3.1, each ad and their displaying page. If at least one category is in common, the ad is classified as contextual. To filter generic (i.e. not customized) ads, we create a number of non-overlapping user profiles (i.e. profiles without any categories in common), and perform 10 requests to the tested pages[3]. Ads that are served independently of the requesting profile are then deemed generic and filtered out.

3.3 User-Profile Reconstruction

Given the targeted ads from the previous step, there are possibly many approaches to infer user information. In our work, we aim at reconstructing the

[3] The number of 10 requests is considered to be enough to get a sufficient ad collection while resisting well to the ad churn [13].

Google-assigned interest categories which are presented as user profiles. In order to reconstruct a user profile, we categorize all of his targeted ads using our Google-based web page categorization tool. The reconstructed profile is then the set of resulting Google categories.

For example, considering the ads provided in table 1, the reconstructed profile will look as follows:

Reconstructed profile
Beauty & Fitness → Fitness
Travel
Travel → Car Rental & Taxi Services
Travel → Hotels & Accommodations
Travel → Tourist Destinations → Beaches & Islandsl
Arts & Entertainment → Entertainment Industry → Film & TV Industry → Film & TV Production
Real Estate → Timeshares & Vacation Properties

Table 2. Profile reconstruction example

4 Experimental Results

In this section, we evaluate the performance of our profile reconstructing technique.

4.1 Experiment Setup

Figure 4 illustrates the setup of our experiments. Our experiments are composed of two main phases:

Profile creation: In this phase, we create a set of profiles corresponding to different web users. Each of these profiles, that we call *targeted profiles,* P_t, is obtained by visiting several websites from a user's real web-history (i.e. list of websites that the user has visited). We refer to these websites as *training sites*. Each of them is visited 15 times to make sure it really affects profiles. We then retrieve the generated Google profile from the Google Ads Preferences page (this phase corresponds to the lower part of figure 4).

Profile re-construction: In this phase, we visit for each targeted profile (P_t) created as described above another set of websites, that we refer to hereafter as *visited websites*. As opposed to the training sites, each visited site is only visited once. The ads are then collected, filtered and the profile reconstructed as described in Section 3. We refer to the set of profiles we obtain as *reconstructed profiles*, P_r (this phase corresponds to the upper part of figure 4).

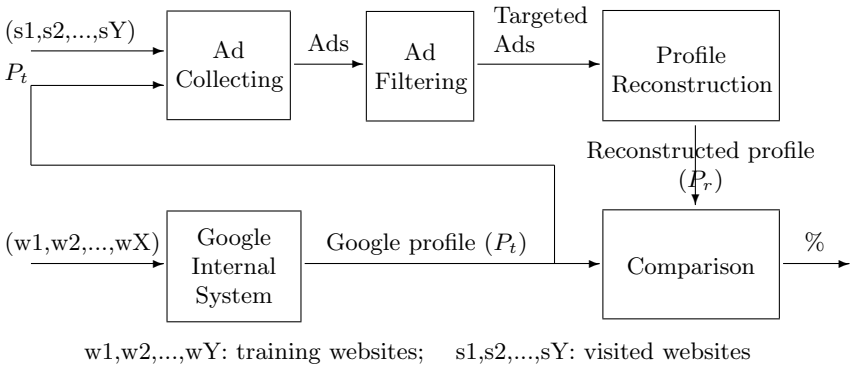

Fig. 4. Filtering targeted ads and inferring user interests

4.2 Evaluation Methodology

Dataset: Our target web-history data comes from a set of 40 volunteers who provided their list of websites they visited during two days. The first X websites in each profile were used as the set of training sites to create P_t. The Y following websites were used to build the reconstructed profiles, P_r, as shown in Figure 5.

In the presented experiments, X was set to 30 and different values of Y were used. The average number of root categories and categories in a targeted profile from X websites is displayed in Table 3.

	# of root categories	# of categories
X = 30	6.64	18.06

Table 3. Profile size statistics

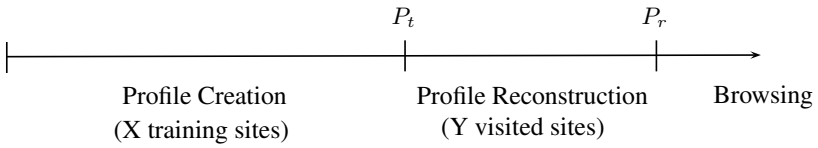

Fig. 5. Profile creation and reconstruction.

Performance evaluation metrics: To evaluate the results, we compare each reconstructed profile with the corresponding original one. We compare profiles using the "same-category", "same-parent" and "same-root" methodologies described in Section 3.1. We evaluate the performance of our technique by computing the average *Precision*, *Recall* and *F-Measure* values of all reconstructed profiles. Precision is the fraction of rebuilt categories that are correct, while Recall

is the fraction of original categories that are correctly rebuilt. F-Measure is the harmonic mean between Precision and Recall, defined as: $F = \frac{2.precision.recall}{precision+recall}$.

In other words, if we denote by $P_{r,c}$ the categories of the reconstructed profile P_r that are correct, and $P_{r,i}$ the categories of P_r that are incorrect, $Precision = \frac{|P_{r,c}|}{|P_r|} = \frac{|P_{r,c}|}{|P_{r,c}+P_{r,i}|}$ and $Recall = \frac{|P_{r,c}|}{|P_t|}$.

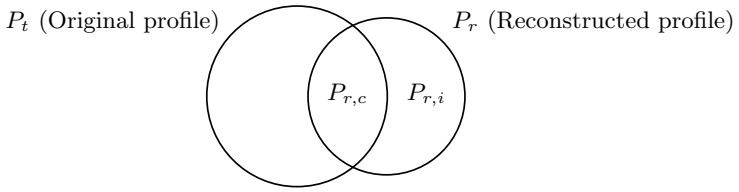

Fig. 6. An illustration of Precision and Recall.

Adversary strategies: In order to evaluate the performance gain obtained by using targeted ads as opposed to only using visited websites, we consider the following three strategies:

- the adversary only uses visited websites ("Sites only").
- the adversary only uses targeted ads ("Ads only").
- the adversary uses visited websites and targeted ads ("Ads & Sites")

Tested scenarios: Finally, we consider two different scenarios, corresponding to two different browsing behaviors:

1. *HotSpot Scenario*: This scenario corresponds to the case where the victim is connecting to an open network and browses the Internet according to his interests. In this scenario, the X training sites and the Y visited sites are related to each others, i.e. generated from the same interest profiles. The goal of this scenario is to show that targeted ads can be used to boost the accuracy of profile reconstruction.
2. *Workplace Scenario*: This scenario corresponds to the case where the victim changes his browsing behavior during the reconstruction phase. In other words, profiles used to generate the training sites and the visited sites are significantly different. This situation happens, for example, when the victim is moving from his home to his work environment. The goal of this scenario is to study how much of the home profile leaks from the targeted ads shown at work.

In the following, we present, for the workplace scenario, how we select the visited websites so that they are largely separated from a user's interests. We first randomly select a set of Google root categories, namely "Autos & Vehicles", "Law & Government", "Books & Literature", "Beauty & Fitness", "Jobs & Education" and "Business & Industrial". We then get for each of these categories

500 websites using the Google Adwords Placement Tool [3]. This tool aims at helping advertisers to select websites to publish their ads. We then get for each user all of his root categories, and select a root category C that does not belong to them. The user's visited sites are then randomly selected from the 500 websites corresponding to category C. For example, if a profile contains 4 root categories: "Law & Government", "Books & Literature", "Beauty & Fitness", "Jobs & Education", then one of the remaining categories, "Autos & Vehicles" or "Business & Industrial", will be chosen for visited websites. We verified that none of our test profiles contains all the six visited categories.

Note that a website classified in a Google category according to Google Adwords Placement Tool may result in another category in Google Ads Preferences. For instance, Google may assign a website W to category "Arts & Entertainment". However, when categorizing this website using Google Ads Preferences, the result may include, in addition to "Arts & Entertainment", another root category, say "Books & Literature". Therefore, we cannot completely guarantee that the visited websites are totally separated from the training ones.

4.3 Result Analysis

Tables 4, 5, 6 and 7 represent the achieved Precision, Recall and F-Measure values in percentage with ($X = 30$, $Y = 10$) and ($X = 30$, $Y = 15$) for the hotspot and workplace scenarios respectively. The rows in these tables specify the category comparison methods used to filter out contextual ads[4]. This comparison method is also used to evaluate the results (i.e. to compare the reconstructed profiles with the original ones)[5]. We remind the reader that these comparison methods are described in Section 3.1. The columns of the table specify the three different cases of profile reconstruction, using "Sites only", "Ads only" and "Ads & Sites", respectively. The tables show that the Ads-based information leakage is significantly high, with precision values ranging from 73 to 82% for reconstructed profiles evaluation based on recovering the root of categories solely from Ads. For example, in case ($X = 30$, $Y = 15$) in the workplace scenario, with "Ads only" and the "Same root" comparison method (used for both filtering and evaluation processes), we achieve Precision, Recall and F-Measure of more than 79%, 58% and 67% respectively (Table 7). The average number of targeted ads we observed accounts for approximately 30% of all collected ads in each case. We note that the results of the row "Same Category" show in general a relatively lower precision and recall values than the results of the "Same Parent" and "Same Root" rows.

Figures 7 and 8 display the variation of Precision, Recall and F-Measure when varying the number Y of visited web sites for each targeted profile, for different comparison methods. We observe that, for a given profile (i.e. when X and therefore $|P_t|$ are fixed), the recall increases noticeably with Y, the number

[4] For example, the row "same parent" displays results when ads are considered contextual if they share the same parent categories with the pages that display them.

[5] For example, the column "same parent" means that two categories are deemed identical if they share the same parent.

	Av.# of targ. ads	Sites only Prec./Recall /F	Ads only Prec./Recall /F	Ads & Sites Prec./Recall /F
Same cat.	14.29	19.66/7.6 /10.96	18.04/7.06 /10.15	18.3/14 /15.86
Same parent	10.94	58.25/29 /38.72	53.67/19.38 /28.48	55.98/42.29 /48.18
Same root	9.24	79.26/51.44 /62.39	73.08/30.06 /42.6	79.6/68.33 /73.54

Table 4. Reconstructing Google profiles performance in Hotspot scenario ($X = 30$ and $Y = 10$)

	Av.# of targ. ads	Sites only Prec./Recall /F	Ads only Prec./Recall /F	Ads & Sites Prec./Recall /F
Same cat.	21.53	19.67/10.28 /13.50	15.71/8.47 /11.01	17.07/17.66 /17.36
Same parent	16.67	54.46/34.44 /42.2	51.26/23.54 /32.26	52.73/50.16 /51.41
Same root	14.4	75.57/61.13 /67.59	82.24/40.3 /54.09	78.5/80.52 /79.5

Table 5. Reconstructing Google profiles performance in Hotspot scenario ($X = 30$ and $Y = 15$)

	Av.# of targ. ads	Sites only Prec./Recall /F	Ads only Prec./Recall /F	Ads & Sites Prec./Recall /F
Same cat.	19.23	2.92/1.05 /1.54	14.73/10.28 /12.11	11.84/10.94 /11.37
Same parent	13.6	9.09/3.99 /5.55	46.31/30.31 /36.64	34.39/31.56 /32.91
Same root	11.2	12.65/6.43 /8.53	78.07/53.96 /63.81	56.49/55.94 /56.21

Table 6. Reconstructing Google profiles performance in Workplace scenario ($X = 30$ and $Y = 10$)

	Av.# of targ. ads	Sites only Prec./Recall /F	Ads only Prec./Recall /F	Ads & Sites Prec./Recall /F
Same cat.	28.11	2.99/1.31 /1.82	13.44/11.95 /12.65	10.89/12.62 /11.69
Same parent	20.3	9.13/5.06 /6.51	44.95/33.8 /38.59	32.75/35.45 /34.05
Same root	17.13	14/8.61 /10.66	79.37/58.12 /67.10	55.85/60.1 /57.9

Table 7. Reconstructing Google profiles performance in Workplace scenario ($X = 30$ and $Y = 15$)

of visited web sites, while the precision is steady. This shows that the number of correctly reconstructed categories, i.e. $|P_{r,c}|$, increases with Y. This result is expected since when Y increases the number of collected ads also increases and as such the amount of available information is higher. However for a given X, the precision is not notably affected by Y, which means that the number of incorrectly reconstructed categories, i.e. $|P_{r,i}|$, also increases with Y.

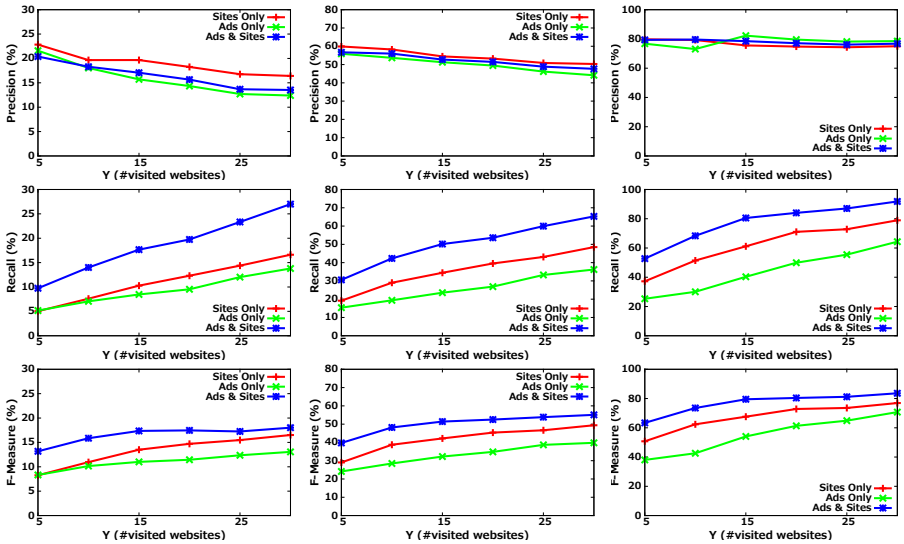

Fig. 7. Precision, Recall and F-Measure with the "Same category", "Same parent" and "Same Root" comparison methods (from left to right respectively) used in both filtering and evaluation processes (**In hotspot scenario with** $X = 30$).

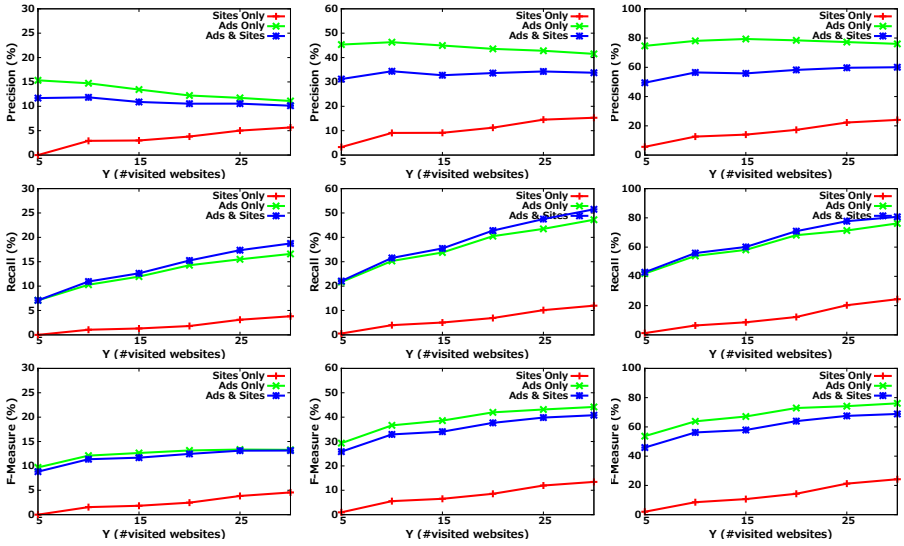

Fig. 8. Precision, Recall and F-Measure with the "Same category", "Same parent" and "Same Root" comparison methods (from left to right respectively) used in both filtering and evaluation processes (**In workplace scenario with** $X = 30$).

In the hotspot scenario, the visited websites are largely relevant to the training websites, therefore reconstructing profiles from "Sites only" achieves the

results as good as, if not better than, the results obtained from "Ads only" (see figure 7). However, when we combine both sites and ads in the reconstruction process, we get nearly the same Precision, while increasing Recalls remarkably (almost the sum of the two cases "Sites only" and "Ads only"). In this scenario, Ads are very useful to boost the performance since they allow the recovery of a larger part of the targeted profiles.

In the workplace scenario, the visited websites are considerably separated from training websites. Therefore, reconstructing profiles from "Sites only" leads to very poor results, whereas the "Ads only" technique achieves significantly better results (see figure 8). By combining sites and ads, we slightly increase the Recall while reducing the Precision. In this scenario, we observe that ads do indeed constitute a "hidden" channel that leaks private information about users.

While the performance of our method when evaluating the recovery of "root" and "parent" categories is notably high, we acknowledge that it can only recover a small proportion of a user's actual categories (Precision varies between 10 and 18% when using the same category method for evaluation, and Recall ranges from 10 to 17%). We believe there are several explanations for this result. First, Google might not be using the user's actual categories to deliver ads, and instead might use the root or parent categories for a broader coverage of users' interests. Furthermore, our ads classification method is probably not optimal and could be improved in many ways. In particular, we took a simple and conservative approach in the filtering step by ignoring location-based ads. In addition, we did not consider remarketing ads that, as discussed in Section 6, may contain additional information about the users recent online activities.

However, even only 10 to 15% of an user's interest categories can constitute a severe privacy breach. To illustrate this statement, we ran our technique on the profile shown in Figure 1, and using targeted ads *only*, we recovered the profile shown in Figure 9. Among the recovered categories, the category "Online Communities → Dating & Personals" may constitute a private piece of information which a user might not be willing to share.

Fig. 9. Reconstructed Profile.

5 Related work

To the best of our knowledge, this work is the first work to quantify the private information leakage from targeted ads content. In the following, we present the most relevant work to our paper:

Privacy-preserving advertising systems. Some initial efforts have been put in designing targeted advertising models yet preventing users from being tracked by ad networks. Among them are Privad [14] and Adnostic [16]. Their main idea is to keep behavioral information at the client side and then to perform the ad selection process locally. The proposed models provide a stronger protection for user privacy than current systems do, but their feasibility is in turn still open to debate. Our work considers a different adversary model. While these schemes try to prevent ad networks from tracking users, we assume that the ad network is trusted and aim at protecting users privacy from eavesdroppers.

Privacy violations using microtargeted ads. Korolova has recently presented attacks that could be used by advertisers to obtain private user information on Facebook [15]. The author showed that an advertiser can manipulate its served ads in order to learn information about users' profiles. This work is complementary to ours, since it considers a different adversary model.

Retrieving user's profile. [11] presented an attack to infer user search history from Google Web search suggestions. While the webhistory webpage is protected by SSL and Google account authentication, the authors showed that a large part of a user's search history can be reconstructed by a simple session hijacking attack.

6 Discussion

Countermeasures. In order to protect against this information leakage, the easiest solution today is to simply opt out of targeted advertising, frequently delete cookies or use ad-blocking software. Initiatives such as NAI (Network Advertising Initiative) [6], DNT (Do Not Track) [1] or TPLs (Tracking Protection Lists) [7] that aim to provide users with tools to restrict tracking and/or behavioral advertising could also mitigate the identified privacy threat. However, these solutions often prevent to target ads or even to serve ads to users.

There exist several possible countermeasures that could be used to target ads to users and mitigate the information leakage identified in this paper. In particular, there are ongoing efforts to design and deploy privacy-preserving ad systems (e.g. Privad [14] and Adnostic [16]) whose main principle is to select ads locally. These solutions make the eavesdropping and filtering of targeted ads, and therefore our inferring attack, much more difficult. Another possible solution would be to send all ad requests and responses (containing DoubleClick cookies and ads content) over secure channels (SSL). However, we believe that this solution needs deeper analysis from the research community and the advertising industry since it might be too costly, not practical or altogether hard to deploy.

Stealing ads preferences via an active attack. The attack presented in this paper is *passive*, i.e. completely transparent to the victim and to the ads

providers. We note that a user's preferences can also be stolen by a simple active attack. If fact, if an adversary is able to steal the victim's DoubleClick cookie, it can connect to his Google Ads preference page and retrieve his preferences. We examined the top 100 commercial websites from Alexa and found that at least 71% of them exchange DoubleClick cookie in clear with remote servers. Stealing a Double Click cookie is then quite easy. We implemented and tested this cookie hijacking attack, and were always able to retrieve the victim's Ads preferences page with a simple request to Google Ads servers. This attack is simple, however as opposed to our scheme, it is active and intrusive.

Remarketing ads. This paper did not consider "remarketing ads", which advertise the services or products of a site that a user has visited. Consider a user who is looking for a hotel in Vigo, Spain and performs some searches on the site `www.hotels.com`. It is very likely that he will consequently receive frequent ads advertising hotels in Vigo while browsing the Internet. Remarketing ads are not only targeting a particular user's interests, but specifically aim to match an exact intention or previous online action. Remarketing ads actually leak much more information about the user. In fact, in our example, they will not only leak that the user is interested in travelling, but also his actual destination i.e. Vigo, Spain. Note that remarketing ads are served independently of Google Ads Preferences profiles. A user will receive remarketing ads even if he empties his ads preferences profile. The only way to stop receiving remarketing ads is to clear his cookies or to completely opt out of targeted ads.

7 Conclusion

In this paper, we showed that targeted ads contain valuable information that allows accurate reconstruction of users' interest profiles. We presented a methodology to categorize and filter targeted ads, which are in turn used to infer users' profiles. Based on both real users' web histories and synthetic users' profiles, we showed that our technique achieves a high accuracy in predicting general topics of users' interests. Additionally, using only a limited number of collected targeted ads we demonstrated that an adversary can capture on average more than half of targeted profiles. The algorithms described in this paper are simple and probably not optimal. We believe they could be improved in many ways.

Many people claim that the main issue in online behavioral advertising is not related to ads personalization itself, which allows users to receive useful ads, but rather to the fact that it requires users' activities tracking. In this paper, we show that ads personalization can also be harmful to users' privacy and does actually leak sensitive information such as users' profiles. We also note that this information leakage is not specific to online behavioral advertising, but in fact exists in any personalized content (news, searches, recommendations, etc.). As the web is moving toward services personalization almost everywhere, special attention should be paid to these privacy threats. This paper contributes in the understanding of possible privacy risks of content personalization.

Acknowledgments

The authors are grateful to numerous colleagues for thought-provoking discussions on an earlier version of this paper, and to the anonymous reviewers for their valuable comments.

References

1. Do Not Track. http://donottrack.us/, 2011.
2. Google AdSense Help. https://www.google.com/adsense/support/bin/answer.py?hl=en&answer=10528, 2011.
3. Google Adwords Placement Tool. http://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=179238/, 2011.
4. Google Display Network. http://www.google.com/ads/displaynetwork/, 2011.
5. Google Privacy Policy. http://www.google.com/intl/en/policies/privacy/, 2011.
6. Network Advertising Initiative. http://www.networkadvertising.org/, 2011.
7. Tracking Protection Lists. www.privacyonline.org.uk/, 2011.
8. Google Ads Preferences. http://www.google.com/ads/preferences/, 2012.
9. Personalized Advertising from Microsoft. http://choice.live.com/AdvertisementChoice/Default.aspx, 2012.
10. Yahoo! Ad Interest Manager. http://info.yahoo.com/privacy/us/yahoo/opt_out/targeting/, 2012.
11. Castelluccia, C., Cristofaro, E., and Perito, D. Private information disclosure from web searches. In *PETS* (2010).
12. Doctorow, C. Scroogled. http://blogoscoped.com/archive/2007-09-17-n72.html, 2007.
13. Guha, S., Cheng, B., and Francis, P. Challenges in measuring online advertising systems. In *Internet Measurement* (2010).
14. Guha, S., Cheng, B., and Francis, P. Privad: Practical privacy in online advertising. In *NSDI* (2011).
15. Korolova, A. Privacy violations using microtargeted ads: A case study. In *ICDM Workshops* (2010).
16. Toubiana, V., Narayanan, A., Boneh, D., Nissenbaum, H., and Barocas, S. Adnostic: Privacy preserving targeted advertising. In *NDSS* (2010).
17. Valentino-Devries, J. What they know about you. The Wall Street Journal, July 31, 2010.