TRAP: using TaRgeted Ads to unveil Google personal Profiles

Mauro Conti
University of Padua,
Padua, Italy
Email: conti@math.unipd.it

Vittoria Cozza, Marinella Petrocchi, Angelo Spognardi
Institute of Informatics and Telematics of CNR
Pisa, Italy
Email: {name.surname}@iit.cnr.it

arXiv:1510.04031v1 [cs.SI] 14 Oct 2015

Abstract—In the last decade, the advertisement market spread significantly in the web and mobile app system. Its effectiveness is also due thanks to the possibility to target the advertisement on the specific interests of the actual user, other than on the content of the website hosting the advertisement. In this scenario, became of great value services that collect and hence can provide information about the browsing user, like Facebook and Google. In this paper, we show how to maliciously exploit the Google Targeted Advertising system to infer personal information in Google user profiles. In particular, the attack we consider is external from Google and relies on combining data from Google AdWords with other data collected from a website of the Google Display Network. We validate the effectiveness of our proposed attack, also discussing possible application scenarios. The result of our research shows a significant practical privacy issue behind such type of targeted advertising service, and call for further investigation and the design of more privacy-aware solutions, possibly without impeding the current business model involved in online advertisement.

I. INTRODUCTION

Online advertisement (also referred as "ad" in the following) generates a business of hundreds of billions dollars.[1] Online ad companies help advertisers to reach the best possible consumers, i.e., those users that would be more sensible to a specific type of ads. Google is currently dominating the online ad market, thanks to two dedicated services: AdSense[2], through which everyone can easily create advertisements.

The Google ads framework involves three main actors: the *advertiser*, the *publisher*, and the *customer*. The customer is every user navigating the Web and seeing online ads. The advertiser wants to reach the customer, in order to show her ads, the publisher sells a space on her website to the advertiser: ads are shown to the visitors of that website. In the whole process, Google acts as a mediator between advertisers and publishers, telling which advertisement has to be shown on which website, depending on the type of visitors that website has. Furthermore, the advertiser can specify the type of

audience to reach, achieving the so called *targeted advertising*. Targeted advertising has the double advantage (when compared to non-targeted ones) to be interesting for the advertiser and—at the same time—pleasant for the customer, which receives ads closer to her interest. Nonetheless, the Google targeted advertising system is exploitable to reveal personal information of the customers: knowing that a specific ad has been shown to a user may simply reveal her interests [1], [2]. *Contributions:* In this paper, we design and test a novel attack to discover personal data (such as users interests, stored in the Google ads users' profile), based on ad impressions. Our approach is particularly novel in the fact that we can reconstruct the user profile "remotely", i.e., without having any direct access to the advertisements the user sees.

Roadmap: In the following section, we present the Google targeted advertising system. Section III describes the attack that can be remotely launched to discover the users' profile. In Section IV, we show our experimental results that confirm the feasibility of the proposal. Section V discusses related work in the area. Section VI gives final remarks.

II. THE GOOGLE TARGETED ADVERTISING SYSTEM

Google offers a complex and rich system for targeted advertising, through the so called DoubleClick infrastructure. This can manage several types of ads: display-based ads (located within pages of websites), search-based ads (which appear among the results of a Google search), ads in the YouTube platform and ads in the Gmail service. In this work, we will focus on display-based ads.

The two main components of the Google targeted advertising system are AdSense and AdWords. AdSense is used by all the publishers that aim at being part of the Google Display Network (GDN), i.e., the network of websites that give their availability to host ads (*ad display* websites). AdWords is the service used by the advertisers to create the ads to be displayed on the ad display websites. Being part of GDN with AdSense has a monetary return for the publishers. Indeed, after a certain number of *impressions*, the publisher obtains an amount of

---

[1]Techcrunch.com: Internet Ad Spend To Reach $121B In 2014, 23% Of $537B Total Ad Spend, Ad Tech Boosts Display: *on.tcrn.ch/l/qsMM*
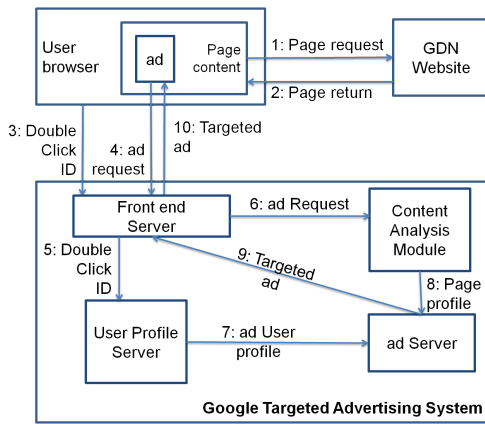
[2]*www.google.com/adwords/*

Figure 1. Targeted Advertising based on user profiles and page profile

Figure 2. Ad user profile for Google ads.

money as a reward. Correspondingly, buying ad spaces with Google AdWords allows the advertiser to generate and manage several type of ads, including those with high personalization. The Google patent on "Targeted Advertising based on user profiles and page profile" [3] describes (at a high level) the way Google provides personalized advertising.

As illustrated in Figure 1, when a user visits a GDN website, she may visualize, along with the page content, both generic and personalized advertisements. Personalization is possible because Google monitors the GDN website content and users' navigation behavior. In particular, for each GDN website, a Content Analysis module evaluates the content of the GDN web pages and builds a *page profile* for each of them, associating to each page a list of *topics*[3]. Topics include general categories (e.g., sport, car, art, music, and movies) as well as sensitive and detailed information (e.g., pathologies and sexual interests). Then, Google builds *ad user profiles*, monitoring and learning users' behaviors when they navigate the GDN websites. When a user visits a GDN website, her browser exchanges a DoubleClick cookie tracing the navigation behavior, as described in the Google patent [3]. The main elements considered are the list of the visited websites, website topics, webpage referral, time spent on each website, number of times the user went to the website, geo-location of the user IP address (if available), visualized advertisements and number of clicks on them. The navigation behavior results in the definition of the user's *interests*[4], which, collectively, represent the "ad user profile". The patent in [3] describes how Google builds both page profiles and ad user profiles.

As shown in Figure 2, at *www.google.com/settings/ads* is possible to access the ad user profile. It is worth noting that Google stores the ad user profile (gender, age, languages, and interests) through the DoubleClick cookie. Interestingly, if the user signs in with her Google account, the ad profile would access gender, age and languages explicitly inserted by the user when creating her account. In Figure 2, the gender and the age are instead inferred by Google according to the navigation
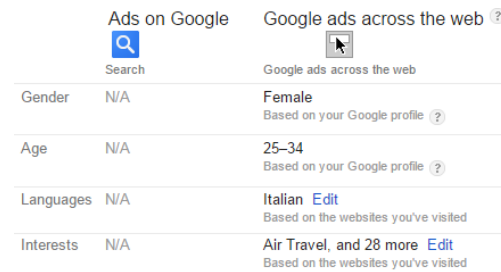
activity of an unlogged user across the Web.

A. Google AdWords: building and targeting a campaign

AdWords makes it easy for an advertiser to create online advertisements by means of a visual browser interface. To configure a campaign, the user starts specifying the "good landing URL", i.e., the destination URL of the user that clicks on the advertisement. Then she has to define the advertisement to show and the keywords describing the good (in case of search-based ads). Afterwards, it sets the *placement*, i.e., the set of websites where the ad will be displayed (in case of display-based ads), and the time and the geographical location the ad will appear. The next steps consist of selecting the *audience* to reach (details on the kind of available audience are discussed below) and in setting the total budget and the *bid* (i.e., how much the advertiser will pay for each ad impression).

The choice of the audience is a very important step, since it determines the type of users that will see the ads. There are several types of audiences: people who previously visited the advertiser website (*in market audience*); people with specific age, gender, parental status and so on (*demographic audience*); and, people interested in the specific products or services of the advertiser (*affinity audience*). An *affinity audience* consists of aggregated users that have demonstrated a qualified interest in a certain topic. In particular, being part of an affinity audience means featuring common navigation behavior on GDN websites with specific topics. Examples of pre-defined affinity audiences are Cooking enthusiasts, Sports fans, Movie lovers, and many others[5]. In addition, AdWords further refines the campaign an advertiser would like to launch, by giving her the possibility to reach exactly the audience she is looking for, specifying a list of websites users should have visited before being exposed to the campaign. Having visited such websites let users be part of *custom affinity audiences*[6].

To represent the relation between topics, interests and audiences, we refer to Figure 3. Mike, John and Leonard belong to the affinity audience "Sports Fans". This is because they share a common navigation behavior through websites related to the same topics (Sports, Soccer, Sports News). Furthermore, checking the users' interests (the ones in the ads
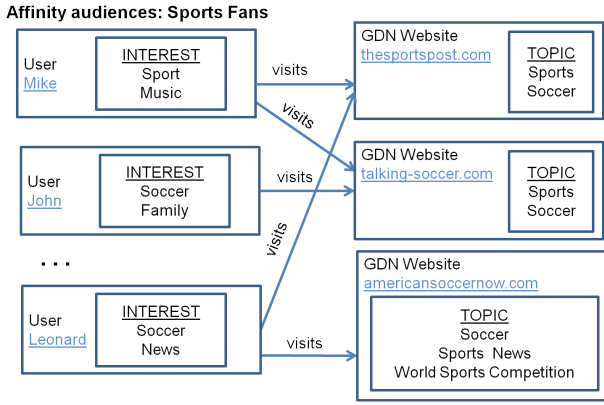

Figure 3. Example of relation between topics, interests and audience

settings shown in Figure 2), we can see that the three users are interested in Sport and Soccer.

Regarding a campaign budget, the advertiser needs to choose the total budget and to set the bidding option for each ad. AdWords provides several bidding options, depending on what matters most to the advertiser and her business. Advertisers can focus on clicks, impressions, and conversions/acquisition, where acquisition means buying a good, booking a reservation or also registering to a mailing list. Available AdWords bidding options are "cost per click" (CPC), "cost per 1000 impressions" (CPM), "cost per acquisition" (CPA) and automatic bidding. Choosing the appropriate bidding is very important, since Google makes online auctions—based on the received bid values—among all the competing advertisers. Then, Google delivers the winning ad to the user when the latter visits the publisher page.

AdWords provides a friendly interface showing how often the ad has got an impression or a click, and if it has reached some demographic or affinity audience.

III. TRAP: TARGETED ADS TO INFER AD PROFILES

As shown in previous work (further discussed in Section V), having access to displayed ads can lead an attacker to know users' personal information, like their interests and their navigation behavior. Here, we go beyond the state of the art by showing settings for and feasibility of a remote attack capable of discovering the user affinity audience and, consequently, inferring the information in the user ads settings. Indeed, since affinity audiences are set according to users navigation habits, if an attacker is able, e.g., to discover that a user is in the affinity audience "Job Seekers", this means that she has visited recruitment and employment agency websites. This can lead to infer that she is looking for a new job and raise privacy concerns (think, for example, if such information is disclosed to the current boss of the user).

TRAP is a system for inferring user affinities of a victim user, whenever a victim visits a website controlled by the attacker, where an advertisement controlled by the attacker is displayed. Correlating the information from the website (such as the log of visiting IP addresses) and the information

collected by the AdWords system, the attacker could reconstruct information related to the (supposedly private) affinity audiences of the victim.

A. Attack setup

To set up all the components, the attacker follows those steps:
1) Website set up. She sets up a website that will be part of the GDN, registers it on AdSense and monitors the visitors' accesses;
2) Ad campaign set up. She sets up an ad campaign through AdWords, with a series of ad-hoc advertisements for several (different) affinity audiences, which will be intentionally displayed only on that particular GDN website;
3) Revealing the affinity audience. She analyzes and compares the AdWords data with the web users' accesses, to infer the affinity audiences and, eventually, the users' interests.

In the following, we detail each of the steps and we describe how we have realized TRAP and the attack.

1) Website set up. The first step consists in setting up a page where we can host advertisements. For this purpose, it is enough to create and manage a website and ask Google to add it into the GDN, in order to serve advertisements and, eventually, to use it as controlled location of our ads.

We have created a fake project website (available at *www.monadsproject.com*). We have populated the website, adding consistent and relevant contents, to make it suitable for serving ads through AdSense, joining the GDN. A website can be part of the GDN only when its content is sufficient enough for the Google crawler to automatically infer its topics. After the AdSense analysis of the page contents, our website has been assigned to many topics, like, e.g., Scientific Institutions, Scripting Languages, and others. We have also added a privacy policy explaining how we manage collected data. The website includes a landing page where the user is explicitly asked to proceed to visit the website and share her data or not, to obtain her agreement before acquiring and analyzing her data. Her preference is stored inside a cookie on her browser.

Finally, to put in place the attack, there is the need to monitor in real-time the visitors activity on the website: this is fundamental to identify the access of the victims to the website, in order to subsequently correlate it to the shown advertisement. For this step, several solutions are possible, like directly reading the server log or relying on professional statistic providers. For our experiments, we relied on StatCounter[7], which, beside the log of each website access, also provides synthetic and aggregated statistics of the visited pages.

2) Ad campaign set up. The second building block consists of setting up a personalized advertisement campaign—intentionally limited to appear on our website. We have registered into AdWords and we have followed standard procedures. To realize the attack, it is important to create a campaign that relies on affinity audiences comprising the interests we would like to extract from the victim's ad profile. Moreover, it is critical to specify that the only website displaying the

---

[7]*www.statcounter.com*

advertisement must be ours: actually, this is the main trick to track the audience, and, consequently, the interests of our victims. As it will be clear in the following, if the ad is displayed in more than one website, this perturbs the statistics on AdWords, making the association victim–interest no more possible.

We have created a Google AdWords campaign named *european_projects*, with an advertisement group *mib_project* that contains our TRAP ad, i.e., our hook for the attack. Each of the possible features of the campaign should be carefully set, keeping in mind the target user, victim of the attack. Beside the choice of the placement website, other elements allowing to restrict the range of users that will receive an impression with our TRAP ad are, e.g., the geographical area, the gender, and the age.

The most important element to configure for the ad campaign is the affinity audience to target. AdWords will take note of the number of impressions of our TRAP ad, for each of the audiences set in the campaign. Being able to link each impression to a specific user means to reveal the affinity audience of that user. In our implementation, we have selected ten affinity audiences (among others: cooking enthusiasts, pet lovers, sports fans, music lover, and health & fitness buffs).

Another important element to be set is the budget, together with the bidding. Among all the advertisements that match their affinity audiences with the user interests, Google chooses the one with the higher bid (for example CPM, cost per 1000 impressions). Consequently, for the attack to be effective, an adequate bid has to be set, in order to guarantee that it will be chosen among all the competing advertisements. If the attacker wants to be really effective, she can set a very high CPM so that, with high chances, her victims will receive the impression of her TRAP ad. We have created an ad with a bidding based on the number of impressions, however the AdWords interface reports both impressions and clicks (Figure 4).

3) Revealing the affinity audience. The third step leads to reveal the affinity audience of the victim. Each time a new user comes onto our website and our TRAP ad receives a new impression, the AdWords increments the counter of the relative affinity audience (Figure 5). Moreover, since the webpages that impress the TRAP ad are hosted on our website, we can access the log or the single visits (as above described) to our page. Then, putting in relation the affinity audience that received a new impression with the visit of the user, we can associate each user to the relative affinity audience. This, in practice, means inferring her navigation behavior and, eventually, her interests. Moreover, with these settings, the TRAP attack can also keep track of the users that clicked on our TRAP ads, not just the impressions. Analyzing the clicks is not considered in this work, but can be used for further analysis.

One limitation of this approach is that the number of impressions and clicks received by ads are updated every 30 minutes. Then, in order to be effective, the attack has to be suitably timed to be sure to make distinguishable the access of the victims. However, this limitation can be easily bypassed if the attacker creates and monitors a single website for each of
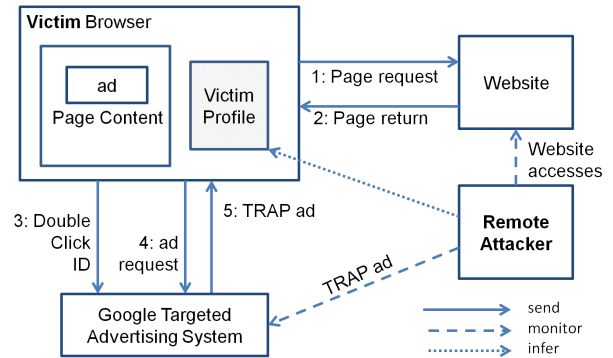

Figure 4. AdWords Impressions

Figure 5. Attack scenario

her victims: even if it can be tedious and time consuming, it is simple and can also be easily automatized, with a random website generator with a database of pages with real contents.

We highlight that the TRAP attack is completely silent and transparent for the victims. It allows to discover users interests by only analyzing the ads Google showed them.

IV. ATTACK VALIDATION

We show the feasibility and the effectiveness of the attack by implementing it in a controlled environment (i.e., not involving real users). In particular, we emulate ten users with ten different ad profiles. We let them visit one or more websites with certain topics (e.g., sport), from a web browser with clean navigation data. As a result, Google adds to the ad profile of each users the related interests (e.g., sport or some subcategory, such as baseball, tennis, etc.). Table I reports the websites visited by the 10 users, with the related websites topics.

We leverage the Google display planner for selecting the

Table I
FILLING AD USER PROFILES

User	Visited Websites	Visited Websites Topics	User's Interests
u1	theatrehistory.com	Arts & Entertainment; Acting & Theater; Literary Classics	Acting & Theater; Broadway & Musical Theater; History
u2	carbuzz.com	Autos & Vehicles; Custom & Performance Vehicles; Vehicle Brands	Autos & Vehicles; Performance Vehicles; Vehicle Brands
u3	delish.com	Cooking & Recipes; Fruits & Vegetables; Food	Beverages; Cooking & Recipes; Fruits & Vegetables
u4	movieinsider.com	Movies; Online Video; Movie Reviews & Previews	Movie Reviews & Previews
u5	www.soundjay.com	Arts & Entertainment; Samples & Sound Libraries; Audio Files Formats & Codecs	Audio Files Formats & Codecs; Samples & Sound Libraries
u6	kingpet.com	Arts & Entertainment; Pets; Dogs	Arts & Entertainment; Contests, Awards & Prizes; Dogs
u7	thesportspost.com	Sports; Team Sports; Sports News	American Football; Baseball; Fantasy Sports; Sport News
u8	onetravel.com	Travel; Air Travel; Travel Agencies & Services	Air Travel; Travel Agencies & Services
u9	talkaboutmarriage.com	Arts & Entertainment, Family & Relationships, Troubled Relationships	Family & Relationships, marriage; Parenting, Childcare
u10	youbeauty.com beautytask.com beautyandtips.com	Beauty & Fitness	Apparel; Bodybuilding; Cosmetology & Beauty Professionals; Hair Care; Make-Up & Cosmetic; Skin & Nail Care

websites to visit in order to fill the profiles as we like. The Google display planner of AdWords allows to select those GDN websites that better match a given topic. The planner shows an ordered list of the websites more relevant to certain interests, specified by selecting a interest category or by simply entering keywords. For each website, it displays a set of charts related to the available visitor population (age, sex, devices), together with the topic of the website and the format of ads they display. As an example, when selecting Sport, the most suitable website is the first online sport magazine *thesportpost.com* . Opening the website details, it is possible to see that it hosts ads in any format and covers the following topics: Sports, Team Sports, Sports News. When a visitor with a clean profile visits this website, Google infers the following interests because of this online activity: American Football, Baseball, Fantasy Sports, Sport News.

A. Experiment and Results

We let the ten users (with ad profiles filled according to a specific past navigation activity) visit the Monads website, which displays the TRAP ads. For each user, we combine information gathered from our monitoring activity with the data taken from AdWords. By doing so, we are able to match a visit event with an impression event and we infer the affinity audience of each user. We ultimately discover the interest of each user since we exactly know which is the TRAP ad each user sees on Monads, because a particular audience (e.g., *Sports Fans*) is incremented by one.

In Table II, we show the affinity audience that has been inferred for all the users under investigation. The result of the experiment confirms that TRAP is able to discover the users' interest through remote management and monitoring of targeted ads. As shown in Table II, in facts, the audience learned with our approach corresponds to the user interests or, at least, to a subclass of them. Many associations are obvious and self-explanatory. User *u10* needs some more explanation, since Fitness is a subcategory of Beauty: Health and Fitness buffs are those that care about their body and appearance.

B. Attack optimization

Aiming at raising the probability to attract the victim to the attacker website, a possible optimization of our approach is

Table II
REAL USERS' INTERESTS *vs* AFFINITY AUDIENCE INFERRED BY IMPRESSION

User	User's Interests	Affinity Audience of the Displayed Ad
u1	Acting & Theater; Broadway & Musical Theater; History	Art & Theater Aficionados
u2	Autos & Vehicles; Performance Vehicles; Vehicle Brands	Auto Enthusiasts
u3	Beverages; Cooking & Recipes; Fruits & Vegetables	Cooking Enthusiasts
u4	Movie Reviews & Previews	Movie Lover
u5	Audio Files Formats & Codecs; Samples & Sound Libraries	Music Lover
u6	Arts & Entertainment; Contests, Awards & Prizes; Dogs	Pet Lover
u7	American Football; Baseball; Fantasy Sports; Sport News	Sports Fans
u8	Air Travel; Travel Agencies & Services	Travel Buffs
u9	Family & Relationships, marriage; Parenting, Childcare	Family focused
u10	Apparel; Bodybuilding; Cosmetology & Beauty Professionals; Hair Care; Make-Up & Cosmetic; Skin & Nail Care	Health & Fitness Buffs

the use of personalized emails and spear phishing [4]. This technique exploits trusted source email addresses (spoofed or not) to invite victim users to visit some website or to pass a unique argument when requesting a webpage (like *http://monadsproject.com?uniquearg=xyz*). Clearly, inserting a tracking element in the email (like *uniquearg=xyz* above), let the attacker infer the victim's interests if she visits the website as suggested in the email. The tracking element could also be a unique embedded image, used to monitor when the recipient accesses the content of the email: this would require a download from the website, unique for the target recipient. A similar technique is used by tracking services, like Streak[8] for Gmail. The same approach can be extended to a group of victims, within a set of known IP addresses—for example, the subnetwork of a public institute or a company. It can also be extended to collect statistics on the interest that occurs more within the group, leveraging TRAP ads targeted for two complementary audiences, e.g., family focused *vs* travel buffs.

V. RELATED WORK

In literature, several works deal with behavioral targeting in Online Ad Systems, highlighting their potential privacy threats [5]. The novelty of the current work is that it presents and maliciously exploits a feature of the Google advertising system, which allows a remote attacker to infer user personal

---

[8]*www.streak.com/email-tracking-in-gmail*

information, as interests and navigation behavior. In the following we review the main research works related to privacy issues and online advertisement.

1) Privacy violations: In [1], Castelluccia *et al.* show how to reconstruct user profiles from targeted ads displayed on the users' browser. They provide a technique to automatically discriminate among different kind of ads shown to the user: generic, page based, location based, profile based. Finally, from profile-based ads they infer users' interests. This work is complementary to ours, since it considers a different adversary model where the attacker is physically behind the computer victim and sees the displayed ads in her browser. Instead, in our scenario, the TRAP ad allows to infer the victim's data without any access to the victim browser.

In [2], Korolova considers privacy violations in Facebook through microtargeted ads. From publicly available information of a Facebook account and with ad-hoc ads, the author shows that it is possible to infer either private information or information that the user had configured as *Only me*, *Friends only*, and *Hide from these people*. More broadly, one can run campaigns in order to infer age or gender distribution of employees of particular companies, estimate the amount of time employees spend on Facebook, the fraction of employees who are interested in job opportunities elsewhere, etc. Similarly, in [6], [7], Cascavilla et al. show several techniques to retrieve supposedly hidden information from Facebook user profile, while they do not rely on advertisements.

Targeted ads are connected with users' navigation behavior. Even if users may have online interests formed and confirmed over a long time period, also short term browsing activity can significantly impact the user's profile and, consequently, change the type of ads that such user sees. In [8], the authors show that publishers can subtly alter the user's profile, in order to make them the target of the most remunerative ads. In [9], Bechmann highlights how profiling is related to privacy violations through a media economics and management perspective. The author analyzes different perspectives on different social media. Facebook focuses on user interconnections and reputation (vertical view) while Google on topic relevance (horizontal view). The work also highlights the implications of profiling for different stakeholders: advertisers, developers and government agencies. Work in [10] analyzes the Levi's company case. The popular brand traces users' navigation behavior through its e-commerce website, it collects data about users and it sends them to third party websites that provide ads, without the user's explicit consent. When dealing with an e-commerce website, the privacy concern is even higher: camouflaging the tracking of consumers can damage the perceived trustworthiness of the brand.

2) Personalization: Background and mechanisms behind the Google search engine and its advertising system are presented in the Google patents [11] and [3]. While the patents describe the big ads picture, they however do not reveal details about the level of personalization in Google search and displayed ads. Relevant personalization in query results and ads would lead to concerns about the "Filter Bubble"

effect, where users are trapped into information bubbles because the search engine algorithms decide such information is relevant for them [12], [13]. In [14], the authors present a measurement methodology for determining the effectiveness of personalized online advertising. It includes a set of guidelines for researchers that wish to study advertising systems and an analysis of the key factors that determine ad targeting on Google and Facebook.

Finally, work in [15] presents a system for personalized web search based interest tree, a classifier able to learn users interests and profiles from browser history. The algorithm then reorders search results, achieving high user satisfaction. The description of the system includes many details about the features considered by search engines in order to determine the users' interests and to combine them in order to understand how each feature is relevant with respect to such interests.

3) Privacy-preserving advertising systems: As analyzed in [16], there exist technologies enabling the delivery of third-party services with affordable privacy risk. Some initial efforts have been put in designing targeted advertising models yet preventing users from being tracked by ad networks, as *Privad* [17] and *Adnostic* [18]. Their main idea is to keep behavioral information at the client side and to select locally the ads. Furthermore, advertisers cannot target sensitive information, like, e.g., health, finances, ethnicity, race, sexual orientation, personal relationships and political activity.

VI. CONCLUSIONS

With behavioral targeted advertising, the user that visits a GDN website can see ads relevant to her real interests, inferred from her browsing patterns. In this paper, we have seen how the technical details that allow to realize this mechanism can be exploited by an attacker to infer users specific interests, without directly interacting with the users themselves. The methodology can be exploited to target specific users by attackers really determined in discovering their interests, clearly violating the privacy of unaware users. We have also described how to extend our TRAP system in order to overcome the slow update limitation of the AdWords system.

While in this paper we did a preliminary assessment of the feasibility of the attack, we plan to further validate it, considering real user profiles (with the user approval).

VII. ACKNOWLEDGMENTS

REFERENCES

[1] C. Castelluccia, M.-A. Kaafar, and M.-D. Tran, "Betrayed by Your Ads!: Reconstructing User Profiles from Targeted Ads," in *Proceedings of Privacy Enhancing Technologies*. Springer-Verlag, 2012, pp. 1–17.

[2] A. Korolova, "Privacy Violations Using Microtargeted Ads: a Case Study," in *Data Mining Workshops (ICDMW)*, 2010, pp. 474–482.

[3] T. Haveliwala, G. Jeh, and S. Kamvar, "Targeted advertisements based on user profiles and page profile," Nov. 27 2012, US Patent 8,321,278.

[4] J. Hong, "The state of phishing attacks," *Communications of the ACM*, vol. 55, no. 1, pp. 74–81, 2012.

[5] C. Castelluccia, "Behavioural Tracking on the Internet: a Technical Perspective," in *European Data Protection: In Good Health?*, 2012, pp. 21–33.

[6] G. Cascavilla, M. Conti, and I. Y. Davide G. Schwartz, "Revealing Censored Information Through Comments and Commenters in Online Social Networks," in *IEEE/ACM Advances in Social Networks Analysis and Mining (ASONAM 2015)*, 2015.

[7] A. Burattin, G. Cascavilla, and M. Conti, "Socialspy: Browsing (supposedly) hidden information in online social networks," *CoRR*, vol. abs/1406.3216, 2014.

[8] W. Meng, X. Xing, A. Sheth, U. Weinsberg, and W. Lee, "Your Online Interests: Pwned! A Pollution Attack Against Targeted Advertising," in *Computer and Communications Security*. ACM, 2014, pp. 129–140.

[9] A. Bechmann, "Internet profiling: the economy of data intraoperability on Facebook and Google," *Journal of media and communication research*, vol. 29, no. 55, p. 19, 2013.

[10] C. Dwyer, "Behavioral Targeting: A Case Study of Consumer Tracking on Levis.com." in *AMCIS*, 2009, p. 460.

[11] S. Lawrence, "Personalization of Web Search," 2005, WO Patent App. PCT/US2004/030,258.

[12] A. Hannak, P. Sapie, D. Lazer, and A. Mislove, "Measuring Personalization of Web Search," *World Wide Web*, pp. 527–538, 2013.

[13] E. Pariser, *The Filter Bubble: What the Internet Is Hiding from You*. The Penguin Group, 2011.

[14] S. Guha, B. Cheng, and P. Francis, "Challenges in Measuring Online Advertising Systems," in *Internet Measurement*, ser. IMC '10. ACM, 2010, pp. 81–87.

[15] A. Nanda, R. Omanwar, and B. Deshpande, "Implicitly Learning a User Interest Profile for Personalization of Web Search Using Collaborative Filtering," in *Web Intelligence (WI) and Intelligent Agent Technologies (IAT) - Volume 02*. IEEE, 2014, pp. 54–62.

[16] J. Mayer and J. Mitchell, "Third-Party Web Tracking: Policy and Technology," in *Security and Privacy (SP)*, 2012, pp. 413–427.

[17] S. Guha, A. Reznichenko, K. Tang, H. Haddadi, and P. Francis, "Serving Ads from localhost for Performance, Privacy, and Profit," in *HotNets*, October 2009.

[18] V. Toubiana, A. Narayanan, D. Boneh, H. Nissenbaum, and S. Barocas, "Adnostic: Privacy preserving targeted advertising," in *Proceedings Network and Distributed System Symposium*, 2010.