DOI 10.1016/j.comcom.2016.12.016

Online Advertising: Analysis of Privacy Threats and Protection Approaches

José Estrada-Jiménez, Javier Parra-Arnau, Ana Rodríguez-Hoyos and Jordi Forné

Abstract—**Online advertising, the pillar of the “free” cont on the Web, has revolutionized the marketing business in rec years by creating a myriad of new opportunities for adverti to reach potential customers. The current advertising m builds upon an intricate infrastructure composed of a var of intermediary entities and technologies whose main aim i deliver personalized ads. For this purpose, a wealth of user d is collected, aggregated, processed and traded behind the sce at an unprecedented rate. Despite the enormous value of on advertising, however, the intrusiveness and ubiquity of these pr tices prompt serious privacy concerns. This article surveys online advertising infrastructure and its supporting technolog and presents a thorough overview of the underlying privacy r and the solutions that may mitigate them. We first analyze threats and potential privacy attackers in this scenario of on advertising. In particular, we examine the main component the advertising infrastructure in terms of tracking capabilit data collection, aggregation level and privacy risk, and overv the tracking and data-sharing technologies employed by th components. Then, we conduct a comprehensive survey of the most relevant privacy mechanisms, and classify and compare them on the basis of their privacy guarantees and impact on the Web.**

Index Terms—**online advertising, Web tracking, user profiling, privacy risks.**

I. Introduction

Selecting and directing information are crucial in every aspect of our modern lives, including areas as diverse as health, leisure and research. In the past, these processes were largely manual, but due to the exponential improvements in computation and sophistication of software, they are becoming increasingly automated.

The industry of online advertising, lavishly illustrated by Google DoubleClick and real-time bidding (RTB), is an example of the ever-growing automation of these processes, and another crucial aspect of our society — to a large extent, the success of most competitive economic activities is dependent on advertising, particularly on the ability to effectively select and direct information to the right potential customers.

Undoubtedly, the advent of the Internet and the Web has created a myriad of new opportunities for advertisers to target billions of people almost effortlessly. However, online advertising is not only ubiquitous. In the early days of the Web, ads were served directly by the publisher (i.e., the page's owner) following a one-size-fits-all approach. But due to the ease with which Web users can be tracked across their page visits, online advertising has also become increasingly personalized. An example of the sophistication of ad personalization is RTB, which enables advertisers to direct ads to the right user and at the right time, by competing in real-time auctions for the impression of their ads [1].

Fig. 1: Word cloud of terms related to online advertising, tracking, user profiling, and privacy solutions in this scenario. We discuss all these terms in this work. The font size of each of them is proportional to the frequency of occurrence in Google search.

Evidently, personalized advertising is the most effective, and hence the most profitable, form of advertising. According to a recent survey, those ads relying on a user's browsing interests ensure conversion rates[1] that double those of untargeted ads [2]. On the other hand, from the publishers' perspective, online advertising is the pillar that sustains the Internet's “free” content and services.

Nevertheless, advertisers and publishers are not the only entities taking part in this business. In fact, there exists an entire infrastructure at the service of both of them, supported by companies like Google, Facebook and Twitter. Enabled by these and hundreds of other ad companies, targeting mechanisms take charge of selecting and directing ads to billions of users everyday, depending on a number of factors such as the page they are visiting; their browsing history; their IP address or parts of it; their operating system; the plug-ins installed and other information related to their Web browser [3], [4], [5]; and obviously the objectives and budgets of all advertisers for displaying their ads.

User information is therefore an asset fundamental to the efficient and effective delivery of advertising, which is not

J. Estrada-Jiménez and A. Rodríguez-Hoyos are with the Departamento de Electrónica, Telecomunicaciones y Redes de Información, Escuela Politécnica Nacional (EPN), Ladrón de Guevara, E11-253 Quito, Ecuador,
E-mail: {jose.estrada,ana.rodriguez}@epn.edu.ec.

J. Parra-Arnau is with the Department of Computer Science and Mathematics, Universitat Rovira i Virgili (URV), E-08034 Tarragona, Spain, E-mail: javier.parra@urv.cat.

J. Forné is with the Department of Telematics Engineering, Universitat Politècnica de Catalunya (UPC), C. Jordi Girona 1-3, E-08034 Barcelona, Spain, E-mail: jforne@entel.upc.edu.

Manuscript prepared July, 2016.

[1] In online marketing terminology, conversion usually means the act of converting Web site visitors into paying customers.

only handed over to the highest bidder, but to many other third parties that are involved in the ad-delivery process. Unfortunately, evident security risks exist for users when personal, sensitive data about their habits are traded in the name of personalized advertising by an infrastructure that operates in the shadows with virtually no oversight [6]. These security risks can be explained in terms of privacy hazards, social sorting, discrimination, malware distribution, fraud and others [7] [8] [9].

Regarding privacy, serious concerns have been raised by the intrusiveness of practices and the increasing invasiveness of digital advertising. According to recent surveys, two out of three Internet users are worried about the fact that their online behavior be scrutinized without their knowledge and consent. Numerous studies in this same line reflect the growing level of ubiquity and abuse of advertising, which is perceived by users as a significant degradation of their browsing experience [10] [11] [12].

In an attempt to mitigate these privacy and security risks, several approaches have been proposed by a heterogeneous group of actors. Research proposals have concentrated on sophisticated mechanisms to anonymize or block the information leaked to third-parties while trying to remain compatible with the current ecosystem. On the other hand, commercial solutions have primarily focused on blocking tracking mechanisms at the cost of seriously damaging the Internet business model.

A. Contribution and Plan of this Paper

This paper presents a "big picture" of the current state-of-the-art of academic and industry solutions that aim at protecting Web users from various privacy threats posed by the online advertising industry. We begin by introducing the main actors of this infrastructure, the interactions among them, and the technologies enabling the delivery of ads. Our survey of online advertising provides the reader with the necessary depth to understand the intricate dynamics of the current advertising ecosystem, and the privacy risks users are exposed to.

To illustrate the risks posed by online advertising, this article conducts a thorough characterization of the capabilities of the components involved in the ad-delivery process, in terms of type and scope of data collection, aggregation level, and, accordingly, privacy threat. This characterization constitutes the first attempt to define an adversary model that systematically classifies and analyzes the elements of the online advertising architecture.

Having identified the privacy risks inherent to online advertising, our second contribution is a comprehensive overview of the protection mechanisms that may cope with such threats. These mechanisms are examined, among other aspects, on the basis of the location of the mechanism employed, the scope of its application and its protection strategy. A significant part of our analysis is devoted to those privacy mechanisms that operate on the user side, since the opacity of online ad platforms has not allowed further research inside. Our review of privacy mechanisms establishes a correspondence between the privacy risks identified in the first part of this work and the proposals, both from academia and industry, that may address them. Finally, we discuss some future research avenues.

We hope that, by systematizing the analysis of privacy risks and protection mechanisms, this article provides privacy designers and researchers with a far-reaching picture of the current state of affairs in online advertising.

The remainder of this work is organized as follows. Sec. II provides the necessary background in online advertising. Then, Sec. III examines the privacy risks inherent to this scenario. Sec. IV conducts a thorough analysis of the most relevant mechanisms to mitigate such risks. In Sec. V, we discuss the various threats identified and the mechanisms that may address them. Finally, conclusions are drawn in Sec. VI.

II. Background

This section examines the modern online advertising infrastructure, providing the reader with the necessary depth to understand the technical contributions of this work. Specifically, we describe the main actors of the advertising ecosystem, the interactions occurring among them, and the technologies involved in the ad-delivery process.

A. The Online Advertising Landscape - From Past to Present

Advertising is commonly linked to commercial activities that involve branding strategies intended to draw the attention of potential customers. The objective of drawing attention is persuading users to buy a product or, generally, spawning brand image. Historically, however, the way potential customers have been contacted by advertisers to apply such strategies has ended up bothering the ones they aimed at attracting [13].

The main problem of classical online advertising has been commonly the very limited media infrastructure by which ads have been distributed to customers. Without enough resources to target users (e.g., TV viewers or newspaper readers), advertisers used to massively flood the available media with ads which very few people were interested in [14]. The flooded message usually "touched" some customers but the strategy was definitely inefficient. Currently, marketing announcements are still sent to an audience that has a huge aggregate size but which is also ultra-fragmented [15] [16]. This is due to the broad range of available media channels (TV channels, websites, etc.) and the volatility of the attention users put on such channels [17].

Despite its shortcomings, online advertising has been a profitable business and proved to be effective in terms of ROI[2], interaction and tracing of potential customers, and reaching an audience [18]. The truth is also that, in the past, audiences were not as fragmented, and the online ecosystem was not as congested as it is currently. As a result, there were more chances for such traditional advertising strategies to be successful.

With the rise of the Internet, the advertising industry has evolved significantly, especially in terms of its capability of reaching potential customers on an individual basis. Modern online advertising takes advantage of recommendation and personalized information systems to tailor advertising campaigns to the interests of Web users [19]. Thus, thanks to

[2] ROI or return on investment is an indicator used to measure the efficiency of an investment.

technologies like RTB, the core of the advertising business is able to show ads to the right person and at the right time, which implies greater effectiveness [20] [21] [9]. Additionally, current online advertising provides more accountability and transparency since the ad companies are encouraged to agree on prices that directly match the effort undertaken by the seller with the benefits received by the buyer. Consequently, in economic terms, advertising services are traded based on the force of demand and supply [5].

Although the online media has transformed the way advertising is conceived, it was not always so. The online environment was originally overwhelmed by confusion where the impact and fulfillment of advertising campaigns were hardly determined objectively [3] [9]. For instance, advertisers had to acquire inventory of spaces available to publish ads without really knowing if such spaces were shown to people interested in the promoted products. Moreover, the lack of resources of the emerging advertising technologies of that time prevented online actors from optimizing the ad-delivery process.

At present, the online advertising landscape is triggered by advertisers, who create the demand, and publishers, who generate the supply. Websites have become the publishers by excellence since the content they offer attracts people whose interests can be revealed from intrinsic interactions with the Web. Moreover, modern methods of online advertising management have incorporated intermediate entities that help advertisers and publishers navigate the web topology in order to connect them together [9]. Such intermediaries, as explained below, are responsible for providing interactive and automatic ad serving that is able to accurately target the intended audience. The targeting strategy implemented by these intermediary entities has directly influenced the ad-personalization accuracy, but also the level of transparency of the process whereby ads are delivered.

Lastly, it is worth stressing that the money produced by online advertising is currently sustaining most of the “free” content on the Web [22]. The money paid by advertisers becomes revenues that are distributed among the different actors of the ecosystem, including the publisher [3].

B. Online Advertising Players

The modern online advertising infrastructure has become certainly complex and dynamic and, although more players can be identified, three components deploy the main roles in this industry. As illustrated in Fig. 2, these components are advertisers, publishers and ad platforms, and their ultimate goal is to display the right ad to the right user [5] [23]. The former two components represent respectively the demand and supply sides of the economic model that governs an online advertising service [9]. The interactions between such players are commonly enabled by an intermediate infrastructure called an ad platform. Finally, users, whose data and requests are the basis of the decisions made for online advertising services, are not directly considered as part of this infrastructure since they do not receive the revenues of such billion-dollar business.

Advertisers are entities that are interested in promoting a brand or product by showing related ads to potential customers. They are willing to pay for displaying their ads [5] [23], and therefore they are the entities that generate the demand of advertising services. Online advertisers are basically aimed at displaying ads on some spaces of the websites (publishers) users visit. Direct agreements may be signed among advertisers and publishers to regulate the online ad service, but these actors commonly get engaged through intermediate platforms, as shown in Fig. 2. Obviously, the use of intermediary entities makes this process more efficient. Thanks to these entities, advertisers may target ads to the intended audience of their marketing campaigns. Also, through modern online advertising mechanisms like RTB, they may participate directly in this targeting process. These capabilities are crucial for advertisers to face the fragmentation of online audiences.

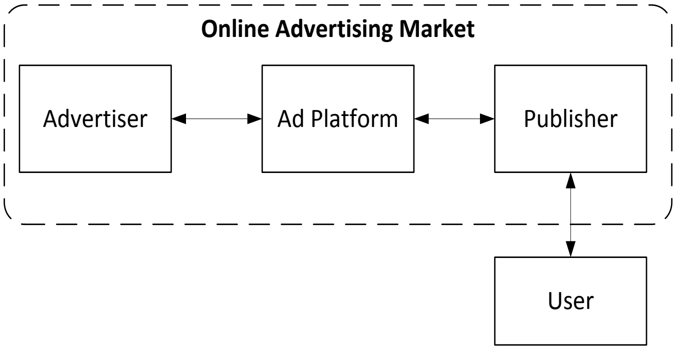

Fig. 2: Main components of the online advertising ecosystem.

A **publisher** is an entity, such as CNN or The New York Times, which provides online content (e.g., newspapers, search engines, blogs, etc.), usually through web pages. Since such content draws the attention of users, advertisers pay publishers to be assigned a space in a website, where they can show ads to a given audience. Commonly, publishers supply advertisers with an inventory of spaces (on their websites) to be filled with marketing messages. Such inventory can be sold by contract or in real time. As depicted in Fig. 3, a publisher is the entity through which a user comes into contact with the online advertising ecosystem.

Ad platforms are groups of entities that connect advertisers with publishers through their demand and supply-side interfaces. In particular, as can be seen in Fig. 3, ad platforms constitute the marketplace where the demand and the supply of online advertising services are matched [5]. In order to effectively reach the currently fragmented online audiences (i.e., a multitude of websites and a pretty scattered attention of users), ad platforms arose to help advertisers and publishers increase the selectivity and efficiency of ad space allocation. Therefore, ad platforms may be considered as the centerpiece of the modern Internet advertising business as they facilitate the matching between the advertising material and users' interests. The accuracy of said matching clearly depends on the ad platforms' ability to track and profile users based on the information that can be mined from their online activity. The ad-targeting process has in recent years become increasingly sophisticated, which has inevitably led to the emergence of numerous agents with very specialized roles. The upshot of this more populated ecosystem (see Fig. 3) is a more automatic,

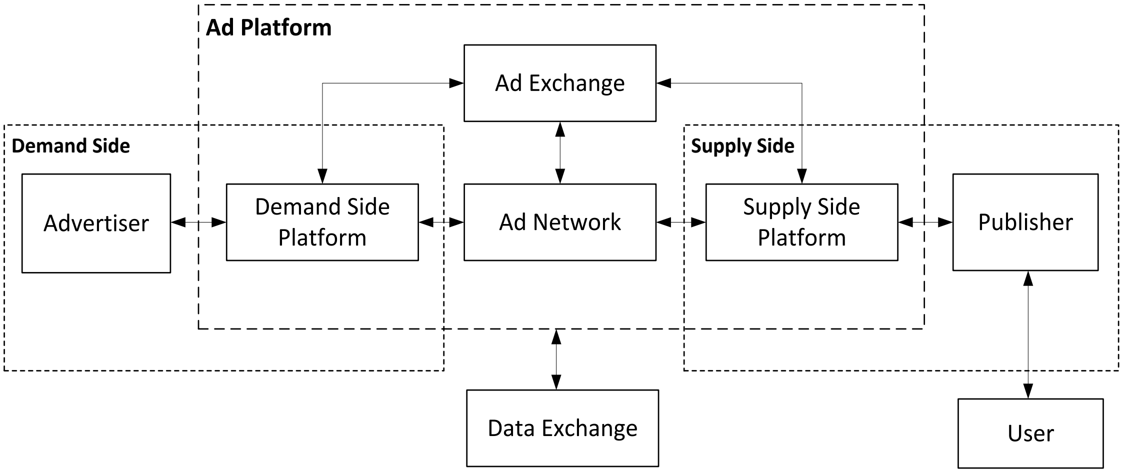

Fig. 3: Disaggregated ad platform scheme and interactions between players.

transparent and flexible ad-delivery process. Throughout this work, we shall refer to ad platforms as *all* the intermediary entities that connect advertisers to publishers.

Originally, ad platforms used to aggregate only the inventory provided by publishers. The aim was to help advertisers get scale and impact (in terms of amount) when distributing their ads; however, scale was not enough. Later, modern ad platforms brought a more transparent infrastructure where advertisers became capable of selecting the users to which they wanted to show ads. To this end, ad platforms integrated certain mechanisms to make the ad-targeting process more accurate, transparent and flexible. Such mechanisms are now implemented by different entities that are part of ad platforms. These entities provide complementary services including aggregation of demand and supply, and optimization of the ad-serving process itself. Some of these entities are *ad networks*, *ad exchanges*, and *demand and supply-side platforms* [9]. Ad networks and ad exchanges are the predecessors of ad platforms. Ad networks began aggregating inventory for advertisers, and ad exchanges evolved to include more dynamic mechanisms to serve ads through automated auctions [24].

Ad networks emerged to help advertisers select and buy ad spaces across the congested and fragmented ad-serving infrastructure. With this aim, such networks used to resell the aggregated ad inventory acquired from publishers to advertisers and related agencies [24]. For those publishers that directly sold their inventory to big advertisers, ad networks became an interesting entity through which to sell their remnant inventory for a good price [3]. Other smaller ad networks were able to give advertisers access to more selective audiences by aggregating more specific inventory from small publishers. Examples of ad networks include GoogleAdSense, Media.net and PulsePoint.

Ad exchanges are ad platforms that currently sell their aggregated inventory of ad spaces by means of auctions. They keep consolidating ad spaces from publishers but offer advertisers and publishers more effective and transparent mechanisms to serve ads [5] [25]. First, ad exchanges place ads based on automated auctions where advertisers “decide” how much to pay for an ad space. The winning bidder is the advertiser that ends up displaying the ad. Secondly, during the auction, ad exchanges share with advertisers contextual information about the user who generates the impression they bid for. Such information helps advertisers decide whether to bid for an ad space and how much to bid for it. The auction is held just after a user requests content from a website partnering with the ad exchange. The whole process may take a few tenths of a second. Theoretically, this yields greater efficiency since the ad-delivery process is distributed among the different components of the ad platform [3]. Part of the aggregation strategy of ad exchanges consists in combining multiple ad networks together. This way, advertisers and publishers are relieved from dealing with so many intermediaries.

Demand-side platforms (DSPs) are entities that work for advertisers, i.e., for the actors generating the demand of ad services. DSPs work on behalf of advertisers, in front of the ad exchange, and help advertisers choose audiences and adequate media to display their ads. By aggregating demand, DSPs are capable of boosting selectiveness and effectiveness for advertisers [3] [5].

Supply-side platforms (SSPs) are entities that work on behalf of publishers, the actors that supply ad spaces to advertisers. SSPs offer publishers an optimized strategy to manage their advertising inventory. Since the task of targeting an ad to a given user involves advanced capabilities and resources, publishers delegate this task to SSPs, with the hope of getting increased demand and profits, despite the congested online ecosystem.

Data aggregators are entities that collect information about Internet users with the aim of profiling their purchasing interests. Data aggregators’ services aim at tailoring ad marketing strategies to the users’ preferences they have learned by means of massive data mining. From data aggregators, another entity called *data exchange* arises. Data exchanges provide demand and supply-side platforms as well as ad exchanges with user data to help them make their targeting decisions.

1) General Operation of Online Advertising: Having shown the main components of the online advertising ecosystem, now we proceed to briefly describe how ads are delivered on the Web.

Currently, ad serving aims at providing automated processes and transparent interactions to advertising entities. However,

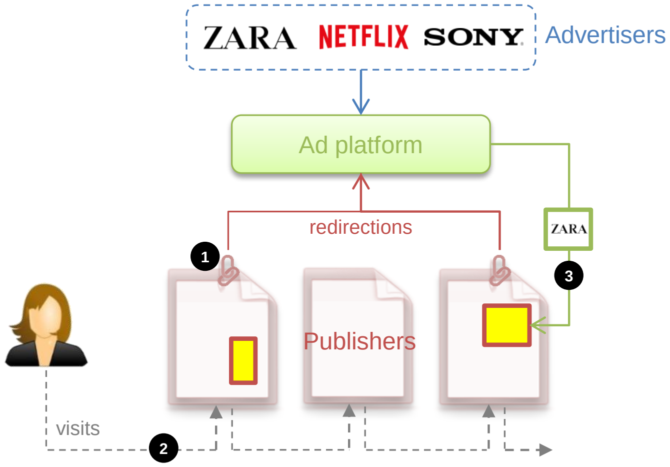

Fig. 4: Current online advertising architecture composed by publishers, ad platforms and advertisers. The ad-delivering process requires that publishers include a link to the ad platform they want to partner with (1); for the sake of simplicity, we consider here a single ad platform. When a user visits pages partnering with this ad platform, the browser is instructed to load the URLs provided by the ad platform. Through the use of third-party cookies and other tracking mechanisms, the ad platform is able to track all these visits and build a browsing profile (2). Based on this profile, the user's location and other parameters, the ad platform uses its targeting algorithm to decide which ad to present on the publisher's page.

there are many interactions involved that make the ad-serving process really complex and completely opaque to the user. In general, when a user visits a website, personalized advertisements are displayed together with the content of the site, as if they were part of the same structure. According to the user's perception, ads seem to be served by the same web server. Although the user participation in the ad-serving process is merely passive, the entire process is triggered by a user's request to download Web content.

This way, when a user's browser sends an HTTP request to a website that is associated with an ad exchange, the website sends back the content the user is requesting. Such content is interpreted by the browser and then displayed to the user. Along with the content, additional code, in the form of ad tags, is sent to the browser and executed automatically. The execution of this code triggers a connection from the browser to the ad exchange in question, which asks for advertisements to fill the ad spaces on the visited page. When the ad exchange receives the ad call, the process of selecting the right ad for the best price is performed by some of the intermediary entities described above. Mechanisms such as RTB and *cookie matching* (CM) are used to ensure the greatest impact on users (which benefits advertisers) together with the highest profits for the ad-serving platform (which includes publishers). Fig. 4 shows the current architecture of online advertising composed mainly by publishers, ad platforms and advertisers, and illustrates the process whereby third-party ads are displayed to users.

C. Supporting Technologies for Ad Serving

The ad-serving process has significantly evolved from the days when advertisers selected the media to deploy ads long before a user visited a website. Currently, advertisers may decide, in real time, which ad to display. As described in the background section, ad platforms take in the order of milliseconds to target an ad to a user based on their preferences and the campaign requirements specified by the advertiser in question. Two main processes are involved. On the one hand, a behavioral profiling task is conducted against a visiting user; this is done on the basis of any information collected about them [20]. On the other hand, automated auctions are used to distribute ads in favor of advertisers, in accordance with their willingness to bid for a particular profiled user.

Mechanisms such as CM and RTB have been developed to support the modern online advertising platforms, by facilitating ad serving personalization and enabling a more efficient and profitable ad distribution system. In the coming subsections, we overview these two mechanisms.

1) Cookie Matching: In order to decide whether and how much to bid for users' impressions, online advertisers require as much information as possible about such users. To come to that decision, the first task of ad platforms is to individuate users so that different attributes can be associated with a (almost) single virtual identity. CM is a mechanism that assists an online advertising platform, and in general a web tracker, in "recognizing" users across the Web. As we shall explain later on, said assistance is key to the bidding processes [26].

CM is based on cookies, which are randomly generated strings of text that web servers send to users' browsers. Cookies are employed to recognize users in subsequent visits. By "identifying" their users, servers are capable of offering personalized services. The same strategy is applied by an ad exchange when serving ads to users, in order to recognize them on a later auction. When a new auction is to be held, an ad exchange sends (ad call) the identifier it keeps about the user to the prospective bidders (advertisers). Such an identifier (cookie) allows advertisers (or their corresponding DSPs) to find any other cookie left on the user's browser in previous auctions. Moreover, an advertiser by itself might have placed cookies on the user's browser from a process unrelated to auctions [3]. Cookies coupled with auction processes may enable advertisers (and other entities) to build profiles of users with information about their browsing history and buying habits.

The process of CM, also called cookie syncing, allows an advertiser and an ad exchange to match the identifiers (cookies) they have about a single user, so that they can share information about them. As stated above, such information enables advertisers to make a more informed decision on whether and how much to bid for an ad impression. A detailed description of how CM works in Google's ad exchange DoubleClick can be found in [26].

2) Real-Time Bidding: Bidding, in general, has represented a breakthrough for the online advertising business. Bidding initially arose for paid-search advertising [27], with the aim of giving transparency to the process of ranking advertisers on search engine results pages. After spamming had affected the quality of search results provided by search engine marketing, and after having realized that such a system prevented smaller companies from participating in the emerging online advertising system, auctions appeared as a mechanism to "democratize" the access to the ad-serving ecosystem [3].

RTB, also called programmatic buying, is an auction-based technology for online advertising. RTB mimics a stock exchange to enable automatic buying and selling of ads [1]. This automatization allows RTB to perform a per-impression bidding just in the moment such an impression is generated. Classic bidding used to take place way before the user accessed the web page where an ad was displayed. Modern bidding, however, is perceived as a real-time process since ad serving is conducted in a fraction of second [28].

RTB enables advertisers to bid for the chance to display an ad on a web page loaded by a user's browser. After such a process, a publisher shows the ad of the advertiser that won the bid. When a user spawns a request from their browser to a website engaged with an ad exchange, a corresponding ad call is generated to the ad exchange. Upon receiving the ad call (asking for advertising), the ad exchange sends a bid request to the advertisers that might be interested in sending ads to a user. Along with the bid request, ad exchanges send valuable information about the user whose impression is being auctioned [29]. Cookies are extensively used by ad exchanges and advertisers to collect and share such information, and thus improve the accuracy of the ad-targeting process [30]. In fact, the very detailed contextual information provided through cookie-related technology helps advertisers and DSPs to make the decision of whether and how much to bid for an impression. After bids are made, a winner is determined during a real-time auction. In a last step, the ad exchange notifies the winner advertiser and its ad is served on the website through the user's browser. This last step may entail a content-delivery network.

III. Privacy Threats in Online Advertising

The pervasive dissemination of online advertising on the Internet and the prevailing need of ad platforms and other intermediary entities to collect a wealth of data about Web users prompt serious concerns regarding user privacy [31] [32]. In fact, much of the concern regarding privacy and thus regarding privacy threats in online advertising are derived from the risks of misuse of this huge amount of user data, which is held by advertising platforms. Said misuse of user information might include common privacy issues such as data leakage, unauthorized collection of data, and sharing with a third-party. Interestingly, as surveyed in Sec. II, the structure of ad platforms and the abilities of their players reflect behaviors strictly coincidental with such privacy issues. In accordance with the above reflection, in this section, we identify the privacy threats specifically inherent to, or arising as a result of, online advertising, based on a characterization of the main players as potential attackers, and of the effects of their capabilities as primary threats.

This analysis and that of Sec. IV exclude the specific context involving mobile devices, albeit much of the following reasoning might still be true for both desktop-based and mobile browsing. Certainly, advertising in mobile communication environments, deserve a separate study, given the complexity of their infrastructures and the growing use of smartphones connected to Internet.

Finally, we want to note that, although the concept of privacy is intimately related to that of information security, the former is addressed here as a particular field of the latter, whose focus is on protecting user data from being revealed, without consent, to potential attackers. Thus, the scenarios in which the user information leaks could be classified as risky.

A. Attacker Model

Privacy criteria are commonly defined in terms of the amount and quality of information that potential attackers might be able to collect about users. Further, characterizing such potential attackers is of special relevance since user privacy is generally measured with respect to the adversary's capabilities as in [33].

Should we consider any entity with access to user data as a privacy attacker, the modern online ecosystem is nowadays plagued by potential adversaries. In the context we address, such adversaries are the multiple intermediate entities developed as part of the online advertising architecture. Although most of these prospective attackers are not directly involved in the raw web traffic spawned by a user, a variety of contextual user information is leaked to ad-serving entities [34] [35]. In general, the information typically collected about a user includes their clickstream, browsing history, shopping habits, preference ratings, entertainment preferences, location, gender, age, and agent string [36].

The online applications and devices (such as browsers and computers) that are daily employed by users lend themselves to the generation of a sort of digital signature that can be subject to fingerprinting. This signature is built with a chain of pieces of information (software installed, plug-ins, and version of applications) that almost uniquely identify a user on the Web. No matter if a user deletes their cookies, they can be tracked online through such a string of data, commonly called an agent string [36].

Even though these items of information might not seem relevant to the identity of a user, several studies have shown that data on some of these "tags" might be sufficient to unambiguously identify a user within a country [37] [38].

Potential attackers in the online advertising ecosystem could be classified as *first* and *third parties*, according to the interaction level of each entity with the user. A first party is directly (consciously) contacted by a user. Nevertheless, third parties are contacted through requests which are not explicitly triggered by users. In this context, publishers may be regarded as the only first-party entities, since the interaction with them is directly made by users; the rest of the components of the advertising architecture depicted in Fig. 3 may be considered as "third-party adversaries". Naturally, the scope of all these potential privacy attackers will vary from local to global according to the amount of users whose information is traded through every component. Of course, such hierarchical scope will determine the aggregation ability and, therefore, the level of privacy risk posed by each of these components.

Publishers can be considered first-party potential attackers within the online advertising ecosystem. Attracting users to its web pages, a publisher receives direct requests from them.

Component	Attacker's role	User collected data	Scope	Aggregation ability level	Privacy risk level
Publisher	First-party	clickstream, local browsing history, preferences, demographics, agent string, identification	Local	Low	Low
Advertiser	Third-party	restricted browsing history, preferences, demographics, identification	Local/Global	Low	Medium
SSP	Third-party	clickstream, restricted browsing history, preferences, demographics, agent string, identification	Global	Medium	High
DSP	Third-party	restricted browsing history, preferences, demographics, identification	Global	Medium	Medium
Ad exchange	Third-party	clickstream, detailed browsing history, preferences, demographics, agent string, identification	Global	High	High
Broadband provider	First-party	every single trace of user interactions with the Web	Global	High	High

TABLE I: Components of our adversary model in the scenario of online advertising.

From such requests, some items of user information can be immediately inferred such as location and agent string. Depending on the type of publisher (news, shopping, social network, rating, etc.), certain information about the user such as gender, age, shopping habits or preference ratings may also be collected. The tracking mechanisms used by publishers are supported on their web log files and first-party cookies.

Advertisers become third-party adversaries since they receive information about users from subtle requests that derive from a user's page visits. Browsing history, location, gender, shopping habits, and other basic contextual data is typically leaked by the online advertising infrastructure so that advertisers can decide whether to bid or not for a given user impression. However, since the described interaction is currently subcontracted to aggregating entities like DSPs and ad networks, the ability of advertisers to directly access user information is significantly diminished.

The ability of **DSPs** to aggregate user information make these intermediaries very powerful potential adversaries to user privacy. Working for thousands of advertisers, a DSP is responsible for selecting the best impressions to bid on. This bidding process is carried out on the basis of both users' metadata and advertisers' specific campaign requirements. Users' contextual data are included in billions of bid requests sent by dozens of associated ad exchanges. Hence, it is difficult to imagine the amount of user information that DSPs are fed with, even without winning auctions. In fact, although ad exchanges recommend not to misuse the contextual information contained in such bid requests, a massive surveillance engine could be deployed through a group of colluding DSPs.

SSPs are the primary source of user information in the current automatic advertising architecture. Helping thousands of publishers interact with other intermediaries such as ad exchanges, SSPs make an offer of an ad space to at least one ad exchange when a user triggers an impression. To give context to such an offer, it is sent along with user data that SSPs gather from different sources. These data may include the visited website, cookies, and browsing information. Thus, SSPs consolidate huge amounts of user data, which raises serious privacy concerns, especially when much of this information comes directly from publishers. From a user's perspective, DSPs and SSPs are third-party adversaries, as they are fed with private, sensitive information that does not come directly from users.

Acting as gateways between buyers (DSPs) and sellers (SSPs), **ad exchanges** are one of the strongest third-party adversaries in our privacy attacker model. These higher-level entities consolidate ad spaces offered by multiple publishers (SSPs) and organize automatic auctions to sell such spaces to advertisers (DSPs). With that objective, ad exchanges concentrate most of the online advertising traffic and the user information used as input to effectively distribute ads. But not only that, ad exchanges also massively distribute such user data to multiple advertisers (mainly DSPs) so that the latter can make their bidding decisions. Given such capabilities of consolidating and indiscriminately distributing user information, ad exchanges are clearly the most powerful privacy attackers of the online advertising ecosystem.

Finally, although they are not strictly part of the online advertising architecture, **broadband providers** are unsurprisingly part of the attacker model we have described. Offering the transport channel that connects every user with the Web, these network-layer intermediaries have privileged access to user information, including that of ad related interactions. Table I summarizes the major conclusions of this subsection.

B. *Classification of Privacy Threats and User Role*

Having specified the adversary model assumed in this work, which we described on the basis of the different intermediary entities involved in the ad-delivery process, next we proceed to classify the corresponding privacy threats based on the capabilities of such entities and the limitations of users.

1) Platform Intrinsic Leaks: The main cause of privacy threats in online advertising is tightly coupled with the infrastructure and capabilities of ad platforms. To start, within this infrastructure, every tracking mechanism is enabled by default; there is not a built-in option for users to disable tracking or ad serving. Additionally, as depicted in Sec. II, this infrastructure is significantly crowded with intermediate entities directly or indirectly fed with user data. Also, it is

Code	Privacy threat	Brief description
T1	First-party tracking	user information leaks out directly from the user side to the publisher
T2	Third-party tracking	user information leaks out from interactions between intermediate advertising entities and the user
T3	Cookie matching	user cookies are mapped and shared between ad exchanges and advertisers
T4	Fingerprinting	an identifying agent string is derived by first and third parties from certain specific characteristics of user applications and devices
T5	Flash cookies	intrusive and persistent cookie technology enabled by Flash-based websites
T6	Canvas fingerprinting	enables user tracking based on a fingerprint generated by the rendering of Canvas HTML5 elements
T7	HTML5 local storage	long persistent cookie-based tracking technology developed as part of the HTML5 language

TABLE II: Summary of the privacy threats examined in our analysis.

Code	User role limitations	Brief description
L1	Lack of awareness	the leakage of personal information is not evident for users in online advertising
L2	Lack of control	user preferences and concerns are not technically enforced by default in online advertising
L3	Bounded technical knowledge	users barely have the technical knowledge to understand and effectively use protection tools

TABLE III: Summary of the user role limitations examined in our analysis.

evident that the business model of online advertising, and so its infrastructure, builds on the collection of as much information about users as possible.

Regarding their capabilities, online advertising platforms carry out practices that support advanced levels of user targeting while neglecting privacy and even supporting the leak of personal data. In this subsection, we briefly examine such practices, which are mainly based on user tracking [35] [39]. Based on the interaction between users and privacy attackers, tracking mechanisms can be classified into first and third-party mechanisms. As we shall see next, these mechanisms mostly employ cookies to individuate users. Table III summarizes these threats.

T1. First-Party Tracking encompasses the activities performed by first-party adversaries (mainly publishers) to collect and analyze user information. Such activities include serving (first-party) cookies directly by the publisher to its users and mining the firsthand information provided by them in their web requests (location and agent string). Depending on the publisher's interaction level with its users, very valuable personal information could be directly gathered by publishers (gender, ratings, social interactions, preferences, shopping habits, health condition). Since the interactions leaking this information are explicitly triggered by the user, they are unlikely to be cataloged as malicious. Thus, detecting or blocking first-party tracking is just as complex, yet the scope of first-party tracking (and thus its privacy risks) is limited to the size of the publisher's audience. Though, some publishers might collude with aggregating entities such as ad exchanges to provide them with aggregated user information [40].

T2. Third-Party Tracking builds on indirect (and non-consented) interactions between intermediate advertising entities (DSPs, SSPs, ad exchanges) and users. Such interactions are generated by content embedded in first-party sites from which user information is also leaked to third parties. The wider scope and higher hierarchy of entities performing third-party tracking for digital advertising facilitate massive aggregation of personal information. However, third party tracking is not only deployed through cookies, but also by means of social plug-ins that may also disclose user browsing information to social networks [41]. Mechanisms aimed at protecting users from privacy risks of online advertising commonly block third-party connections after classifying them as undesired [42].

T3. Cookie Matching is a technology that supports the sharing of user data. Served both by first and third-party adversaries, cookies are the basic tracking technology used in online advertising. Within online advertising, cookies have given rise to concerns about the privacy of users for two main reasons. First, cookies are currently being used to store personal information (such as e-mail addresses), not only identifiers to recognize a user in future visits [43]. Secondly, they enable massive sharing of such personal data through a more refined tracking technology, CM. CM enables an ad exchange to share users' cookie information with multiple potential advertisers so that they can infer contextual user data by mapping their own cookies (obtained from previous interactions with a user) with the ones obtained from the ad exchange [30].

Experiments done by Bashir et al. in [40] report about the ubiquity of CM on today's Web and on how shared information supports highly targeted advertising. It is worth noting that, although using cookies is an old practice originally built upon pretty small pieces of identifying information, they have significantly evolved to become large capacity structures, very popular tracking mechanisms, and increasingly more difficult to delete, as illustrated in Tables IV and V. Accordingly, a great deal of recent research has been done regarding online tracking [44] [45] [46], studied in desktop browsing contexts where the most evolved forms of cookies [47] [48] are subject to analysis.

T4. Fingerprinting, not built on cookies, is also available to support personalized online advertising. It consists in detecting the agent string of users' devices or applications. Thus, no matter if a user deletes her cookies, they can always be tracked online through such an agent string [36]. As a matter of fact, some variations of fingerprinting are commonly used to respawn cookies after a user deleted them. Mayer and Mitchel synthesize in [25] a list of non-cookie web tracking technologies used both from first and third-party entities.

T5. Flash Cookies [47] pose an alternative tracking technology for advertising entities trying to face the advent of mechanisms to block traditional tracking. Flash cookies are more effective in tracking users than common HTTP cookies. In fact, Flash cookies are considered prominently intrusive due to their persistence characteristics (more storage capacity, browser independent storage, and non-default expi-

	Max. storage size	Level of persistence	Storage location	Difficulty to delete	Usage level	Installation	Access level
HTTP cookies	4 KB	low	within the browser	low	remaining	native	one browser
Flash cookies	100 KB	medium	outside the browser	high	declining	through a plug-in	multiple browsers
HTML 5 cookies	5 MB	high	within the browser	high	increasing	native	one browser

TABLE IV: Comparison of the types of cookies that are typically used to track users.

ration) [47] [48] [49]. After online advertisers were accused of misusing Flash cookies (by enabling restoring of deleted HTTP cookies), a study by McDonald and Cranor [50] found that the practice of respawning erased cookies had become significantly less aggressive.

T6. Canvas Fingerprinting is another persistent web tracking technology currently used by some online advertising agents, especially data aggregators [51]. Canvas fingerprinting facilitates tracking by generating a fingerprint of a user's browser from an HTML 5 Canvas element [44]. Such an element might be used by an (first or third-party) adversary to dynamically display, even invisible, text or images in the user's browser. Since the rendering of the Canvas element will slightly vary depending on the web browser's image processing resources, such particular displaying parameters could be used to get a fingerprint that might uniquely identify a user surfing a web page; to do it, certain browser properties are collected such as the list of installed plug-ins [36]. A few first and third-party providers of Canvas fingerprinting have been found from previous studies [44] and the tracking mechanism can be blocked if the provider's domain is known.

T7. HTML5 Local Storage is an even more persistent cookie-based tracking technology, developed as part of the HTML5 web language. Local storage enables more universal user tracking [52] that does not depend on the browser used, does not expire, and offers even more storage capacity, by default, than HTTP and Flash cookies (see Table IV). Such a feature might let some first or third parties store data (within the user's browser) that cannot be deleted when erasing browser's cookies. However, such intrusive tracking mechanisms might be aggressively tackled with lawsuits, especially when accomplished by advertisers, as Wired reported in 2010 [53]. Said misusing of cookies was reported by Hoofnagle et al. in 2012 [52] when they found that some companies had been using HTML5 and Flash cookies to respawn HTTP cookies that had been previously deleted by users. In Table V we summarize some of the characteristics of these tracking mechanisms including their effectiveness in individuating users, and whether the companies using them have faced lawsuits due to the intrusiveness of these mechanisms.

Other intrinsic properties of ad platforms make them pretty susceptible to privacy leaks. For example, the subtlety of their background processes isolates users in a separate dimension where they are unaware of the implicit risks. In addition, as recently reported in [43], relevant user information might be being conveyed in the clear text during real-time auctions. In the same, [29] and [40] reported cooperation between relevant entities such as ad exchanges and publishers, and quantified the derived leakage of users' browsing information. On a last note, chances are that the context information that feeds auctions will reach entities not really involved in bidding processes (or deliberately bidding to lose). Should ad platforms cannot detect such behavior, a cheap massive surveillance tool could be built on top of advertising infrastructures.

2) User Role Limitations: User capabilities are, by default, pretty limited online. Although their interactions fuel ad delivery services, users are unaware of the transactions that are made in the background when they are served an ad, which also reduces their chances to protect themselves. This blindness and lack of control of users is the source of important privacy threats, especially in online advertising systems, where ad services are inherent to web browsing.

L1. Lack of awareness. Historically, online privacy has been a concern for users, as reflected in [54]. However, as explained by Ackerman et al., when faced with an abstract context where the leakage of personal information is not evident (as it might be within social networks), users' concerns get significantly lightened. This attitude of users towards privacy, particularly in advertising environments, is illustrated in [55], which report that users are more concerned about being shown embarrassing ads than about being tracked.

In accordance with said lack of awareness, users hardly notice the relative value of their data within commercial contexts. Evidence on the dichotomy on how users and ad services value user data is offered in [56] and [29], respectively.

L2. Lack of control. In the opaque scenario of online advertising, users cannot protect their privacy adequately. Neither their interests nor concerns can be enforced because users are, by default, passive entities in the advertising ecosystem.

L3. Bounded technical knowledge. Users face an important cognitive barrier that seriously limits their capabilities to manage their protection against privacy threats in online advertising. Even being aware of the risks posed in this context, and having the control to at least mitigate some of them, most users do not have the technical knowledge to understand the logic of protecting themselves within such a complex scenario.

Consequently, in online advertising contexts — unlike what happens in other online scenarios —, leaks of user data are not driven by user explicit flaws but arise from the complex structure and operation of the ad-serving process. Ironically, online advertising was said to offer users more control over advertising exposure than traditional advertising [57].

	Effectiveness individuating users	Ad companies involved	Have led to lawsuits?	Easily erasable from browser?	Usage level	Are intrusive?
HTTP cookies	High	All [44]	No	Yes	Extended	No
Flash cookies	High	`hulu.com`, `about.com`, `aol.com`, Clearspring, Interclick, Quantcast [25] [44]	Yes	No	Extended	Yes
Canvas fingerprinting	Low	Addthis [44]	Yes	No	Limited	Yes
HTML5 local storage	High	Ringleader Digital, Bluecava [25] [48]	Yes	No	Growing	Yes

TABLE V: Tracking mechanisms used in modern online advertising.

C. Impact of Online Advertising Practices on Privacy

Since ad personalization (e.g., based on location, context and interests) increases conversion rates, users' browsing data have inevitably become an asset that nowadays is exchanged throughout the entire online advertising infrastructure [43]. The need to further scrutinize this information to profile and segment users raises serious privacy concerns with respect to social sorting and discrimination, particularly as potentially sensitive information can be inferred from the profile of a reidentified user, such as income level, health issues or political preferences.

Modern auction-based ad delivery requires that processes be executed in real-time, which implies that vast amounts of user information be mined at very high rates. This urgent need might naturally discourage the online actors from protecting user information against privacy attacks. Besides the urgency in which data must be handled, the need to offer tailored ads compels the advertising ecosystem to collect a wide range of metadata. For this reason, practices such as cooperation (collusion) among advertising entities and aggregation are enabled to facilitate massive and often uncontrolled sharing of said information [30]. Since the shared data (sometimes including even the prices paid by advertisers) are not always encrypted, other adversaries, such as Internet providers, come into the picture.

As described in previous sections, online advertising builds on non-transparent interactions among a myriad of intermediary ad companies, which have the ability to profile Web users. As a result, not even publishers are aware of which information is collected and how it is used. In fact, publishers are unaware of what ads are shown to their visiting users. The ad-delivery process involves so many intermediary companies that it is impossible for an ad exchange to control the use of user data by such companies. In fact, cases are known where attackers took advantage of advertising channels to distribute malicious code to millions of users [8]. This lack of transparency obviously prevents users from actively getting involved in the protection of their privacy. Though there are informed users who use transparency and protection tools while browsing, advanced mechanisms are currently implemented by the online advertising ecosystem to counteract cookie removal or ad blocking.

Finally, due to the auction-based policies of the advertising ecosystem, certain users invariably become more economically valuable than others. For example, Olejnik et al. found in [29] that, in terms of prices paid during online auctions, visitors of websites belonging to particular categories are much more relevant than visitors of websites of other categories. Yet, other criteria such as the user location and time of visit might also be used to determine the relevance of the corresponding profiles. Such more relevant users stand out from the rest and gradually their profiles become more identifiable and, as a result, less private. Unfortunately, evidence has been found suggesting that negative discrimination (such as racism) might be performed in online ad delivery [58].

IV. Analysis of Privacy-Protecting Approaches

The privacy risks posed by the tracking and profiling practices of the online advertising industry have motivated a variety of privacy-protecting approaches from academia. These research initiatives mostly rely on mechanisms that may support or complement the current economic model of the Web, while others suggest moderate blocking of third-party tracking[3] to protect user privacy. Other plug-and-play proposals are also available to users and are supported commercially. In essence, such approaches provide users with transparency and control functionalities over their browsing data, yet putting at risk the Web economic model, currently built on the revenues of online advertising, through radical blocking mechanisms.

In this section, we address the main parameters that characterize the current privacy protection approaches in online advertising, in particular, their location, scope of application and strategy. Afterwards, we analyze the most relevant research work and industry proposals which tackle the problem of privacy protection in online advertising.

A. Protection Parameters

Our analysis of privacy mechanisms examines three main aspects, which we proceed to describe.

1) Location: According to the location where the protection mechanism takes place, the current research proposals and commercial solutions can be classified roughly into local and third-party. On the one hand, local mechanisms commonly lie on the user side, for example, in the form of an application running on the user's browser, or as a local service operating in the user's network [61]. Some academic approaches propose migrating the profiling processes required for ad targeting to the user side [35]. On the other hand, third-party mechanisms are implemented with the help of a broker entity, whose location is remote from the user side, and whose aim is

[3] The vast majority of ads today are served by third-party entities [59], [60].

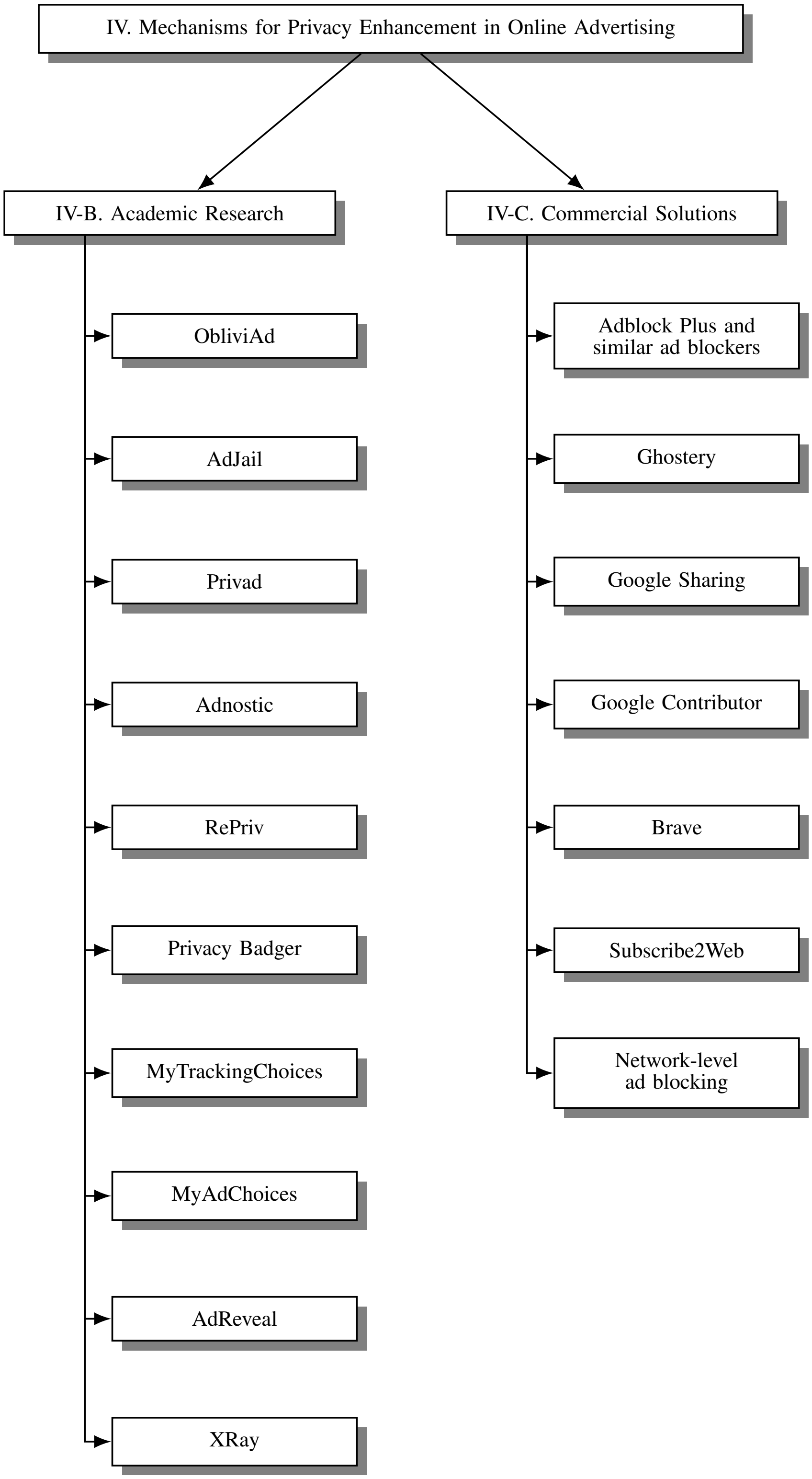

Fig. 5: List of privacy mechanisms, specifically intended for online advertising, that we examine in Sec. IV.

commonly to provide security services such as secure storage of data, anonymization and even user profiling [62]. We would like to stress that, even in the case of broker-based mechanisms, a local application on the user side is frequently required to engage users to said broker.

2) Scope: Depending on the scope of application of the mechanism in question, we may characterize it as local or global. Protection approaches whose scope is local usually aim at adapting a protection mechanism to the structure of the current advertising ecosystem. Hence, the scope of protection offered is also limited to the information and interactions available to the user. On the other hand, those protection approaches with a global scope come in hand with new ad delivery models, pretending to radically change the manner in which ad serving processes currently function, especially with regard to their relationship with users. The majority of these approaches has been envisioned as privacy-by-design models of advertising which would provide users with significant control over their interactions with ad platforms.

3) Strategy: In our classification of privacy technologies, we also consider the principle or *strategy* that rely on. We contemplate five strategies which range from user lack of awareness through *transparency*, to undesired interactions with third-parties by means of *blocking*, *obfuscation*, and *sandboxing*, and to a by-default exclusion of users from the advertising logic through more *inclusive* techniques. Next, we describe these strategies.

Transparency: Undoubtedly, a first step towards privacy protection may be *transparency*. Transparency in this context means allowing users to learn what is going on with regard to their activity and data in online advertising systems. Some of the approaches examined in the coming subsections provide transparency usually by making users aware of the tracking activities behind the scenes, and by allowing them to know how their browsing traces might have been exploited to deliver targeted ads.

Blocking: Blocking is also a very common, although usually radical, strategy of privacy protection in online advertising [42]. Typically, blocking tools inhibit most of the known tracking mechanisms (and thus of advertising) from the user side, or a third-party located on their network. Because the vast majority of ads are delivered nowadays through third-party trackers, cutting of third-party tracking implies eliminating nearly all ads. Originally, blocking mechanisms had been designed as a binary choice, namely, either blocking or allowing all tracking and hence advertising. Nevertheless, recent academic proposals tend to lighten this radical strategy by providing fine-grained control over tracking, by enforcing users' preferences and by using smart and dynamic learning mechanisms [63] [64].

Obfuscation: It consists in perturbing sensitive data in order to preclude an adversary from discovering the identity of its owner and/or deriving private information about them [65]. In the context of online advertising, some privacy protection approaches implement obfuscation by mixing data and metadata of a group of user profiles so that the intrinsic features of individual profiles cannot be recognized. Other approaches build on external brokers to anonymize user data by randomly masking potentially identifying attributes such as IP addresses and cookies.

Sandboxing: Sandboxing addresses security threats by isolating suspicious applications from the resources they rely on. Within online advertising, sandboxing is applied by keeping apart certain critical processes which may give advertising brokers access to sensitive user data. A typical sandboxing example leverages on the execution user profiling on the premises of the user, rather than on the ad-platform side [35] [66] [67].

User Inclusion: With the aim of balancing the Internet's dominant business model and user privacy, some proposals envision a more user-driven ecosystem. In general, giving users more control over their interactions with ad platforms might help achieve said balance. A practical step towards this consists in adapting the protection mechanisms to the needs of users. In this line, most ad blocking solutions have recently started to offer users some personalization features such as blocking per domain and per tracker. Other strategies include the enforcement of user choices over third-party, cookie-based tracking, and the provision of direct interfaces with the advertising ecosystem [66], [68].

B. Academic Initiatives

This section examines in depth the most relevant approaches in the academic literature of privacy-enhancing technologies for online advertising. Fig. 5 provides an overview of the technologies explored in this section.

1) ObliviAd: Proposed by Backes et al. [62], ObliviAd relies on adapting secure-coprocessor-based brokers to the current online advertising ecosystem. The aim of such coprocessors is providing private information retrieval of user data during the delivery of ads to users and the billing to advertisers. To do so, this approach provides different services such as the secure storage of sensitive user data; the encryption of profile information when it is conveyed to the broker side; the encryption of ad information to be displayed on the user side; and finally the obfuscation of billing data to charge advertisers.

While all these services that integrate ObliviAd may offer strong security guarantees through hardware and heavy cryptographic techniques, this is undoubtedly at the cost of a significant increase in complexity and deployment. It is worth stressing as well that network and browser identifiers such as the user's IP address and user agents might still leak, which means that this approach might not be useful against the fingerprinting techniques described in the background section of this work.

2) AdJail: Ter Louw et al. [69] proposes a tool that aims at empowering publishers to isolate the content elements to which ads will have access to. Specifically, this approach allows safeguarding a user's scope and that of the web application by creating a sandbox where ads are executed. From this sandbox, ads may have access to user or publisher content through a configurable set of enforcing policies. Although the aim of AdJail is to protect the confidentiality and integrity of user and publisher data, user privacy can also be provided by applying those policies based on the privacy agreement negotiated between publishers and their users.

The problem of AdJail, however, is that its scope is limited to the publisher's domain. In other words, users can utilize this sandboxing approach only if this mechanism is deployed in the website. In addition, AdJail focuses more on other security services such as integrity and confidentiality, and does not tackle the privacy threats identified in Sec. III.

3) Privad: S. Guha et al. [67] seeks a more private online advertising system and offers to this end an alternative private solution that may adapt seamlessly to the current advertising business model. The authors argue that Privad, their solution, would preserve privacy by keeping a user's browsing profile within a local user application. Nonetheless, they also claim that some information (related to the user's interests and to the ads the they have viewed or clicked) "necessarily" would leave the user's domain.

Protection mechanism	Location	Scope	Protection strategy				
			Transparency	Blocking	Obfuscation	Sandboxing	User inclusion
ObliviAd	local, third-party	local			✓	✓	
AdJail	local, third-party	(1)		✓		✓	
Privad	local, third-party	local			✓	✓	
Adnostic	local	local, global (2)			✓	✓	✓
RePriv	local	local		✓		✓	✓
Privacy Badger	local	local	✓	✓			✓
MyTrackingChoices	local	local	✓	✓			✓
MyAdChoices	local	local	✓	✓			✓
AdReveal	local	local	✓				
XRay	local	local	✓				

TABLE VI: The academic proposals for privacy enhancement in online advertising are classified on the basis of the protection parameters described in Sec. IV-C. (1) AdJail's scope focuses on the publisher; accordingly, we categorize this proposal neither as local nor as global; (2) although Adnostic's protection mechanism lies in the user side, ad platforms would need to adapt to support it.

Privad also incorporates a third-party anonymizing proxy. This proxy would receive the released (and ciphered) user information and, after hiding the user network address, it would deliver this information to an ad platform. Advertisers aiming at delivering ads feed the ad platform with their ads, including information of the profile to which each ad is targeted; and then this information is employed by the ad platform to tailor ads to those profiles. Consequently, this approach uses the proxy to anonymize user information so that the ad platform in question is not able to individuate a user from the preferences reports generated by their browsing activity. Unfortunately, anonymizing strategies like this have proved to be weak [38], especially when demographic information about users is still available for a potential attacker, and when such information is managed by a third-party entity over which a user might not have any control (such as an Internet service provider).

4) Adnostic: It is an academic proposal by Toubiana et al. [70] that implements a more friendly architecture to display personalized advertising without compromising user privacy. Such architecture does not rely on blocking ads but on performing the whole user profiling process within the user domain, so that no personal information is leaked out to third parties.

The ads to be shown to a user are chosen on their side, according to a locally estimated browsing profile. This profile is constructed by processing the user's queries and the content of visited pages. Then, this information is classified within the browser by means of natural language processing techniques. The ads, which are part of a previously downloaded set, are displayed according to the user's interests.

Because personalization is not directly controlled by ad platforms, there are less incentives for advertisers to bid more money to place ads. However, we may expect worse personalization performance since this process takes place on the user side, based only on their browsing data. This is in contrast to the current ad-targeting algorithms implemented by ad platforms which rely on massive amounts of aggregated user data.

In terms of impact on the current infrastructure, on the other hand, Adnostic would eliminate the requirement of intermediary ad platforms, but unfortunately at the expense of less effective ad-targeting. As a matter of fact, the more components of the online advertising architecture are embeded on the user side, the more control the user may have over advertising. Obviously, this would mitigate many of the privacy risks analyzed in Sec. III.

5) RePriv: It is a proposal by Fredrikson and Livshits [66] that aims at carrying out a selective disclosure of user data through a browser-based tool. First, as with the extensions described above, the proposed system would rely on the ability of the browser to capture all the information spawned while browsing the Web; this is the basis for local user profiling. Next, the system contemplates that the interests derived from such user profile are released to third-parties only if the user gives permission. Detailed information about their browsing habits, though, would not be released by default. Finally, the proposed system considers additional modules that would interface with third-party applications interested in having access to user data.

The privacy-preserving strategy of RePriv consists in profiling users locally, so that they have control over the information that is disclosed to ad companies. However, although users are in control of said disclosure, external entities might be collecting such data anyway. Even though at first RePriv might seem an interesting approach, its success in protecting user privacy certainly depends on the disclosure control given to users. Again, such a control may tend to be absolute (as in ad blockers) or could be softened to balance the trade-off between user privacy and the Web business model.

6) Privacy Badger: Much of its functionality was incorporated from an older project called ShareMeNot which was originally presented in [71] by Roesner et al. Currently supported by the Electronic Frontier Foundation (EFF), Privacy Badger is an open-source browser extension developed for Chrome and Firefox [63]. The extension was not conceived as an ad blocker, but as a privacy tool that may prevent non-consented tracking.

The operation of this browser plug-in does not rely on blocking all tracking by default and on static filtering lists (see Sec. IV-C on ad blockers). Instead, it capitalizes on an algorithm to detect and then prevent non-consensual tracking activities. Since the blocking mechanism is not based on the subscription to a deliberate filtering list but on rigorous algorithmic methods and policies, engagements with advertising companies to include blocking exceptions are in principle less likely to occur.

With regard to its graphical interface, this extension looks very similar to Disconnect and Ghostery. The user is shown the tracking companies following their visit to a page. As mentioned above, this tool does not block a tracker unless its algorithm checks it is following the user without their consent. Nevertheless, conducting this checking may take some time. On the other hand, as with most ad blockers, users may individually block or allow the detected trackers, or block only the corresponding tracking cookies. Additional options include disabling the extension on a per-site basis and manually adding a whitelisted trackers domain.

Privacy Badger represents a promising approach to balance the trade-off inherent in online advertising between user privacy and the Web economic model. In fact, besides blocking non-consensual tracking, its developers offer ad companies the opportunity to be whitelisted if they formally promise to respect opt-out mechanisms (e.g., Do Not Track headers), conforming with users' privacy policies [63].

7) MyTrackingChoices: Achara et al. [68] propose a browser extension available for Google Chrome and Mozilla Firefox. The plug-in targets users who are not in general against advertising and accept the trade-off that comes with the "free" content. However, for privacy concerns, they wish to exert fine-grained control over tracking.

This academic proposal relies on the assumption that some categories of web pages (e.g., related to health or religion) are more privacy-sensitive to users than others (e.g., about education or science). Based on this idea, the plug-in allows users to specify the categories of web pages that are privacy-sensitive to them and block the trackers present on such web pages only. As tracking is prevented by blocking network connections of third-party domains, MyTrackingChoices avoids not only tracking but also third-party ads.

The detection of the tracking companies does not rely on existing blacklists, unlike most ad blockers and anti-trackers. Rather, MyTrackingChoices keeps a local list that is built from the pages browsed by the user. This list is smaller and easier to maintain than the list of tracking and advertising domains currently used by Adblock Plus. To decide if a third-party domain is a tracker or not, the tool checks it is present on three or more different domains that a user visited in the past. Since users continue receiving ads on those web pages which belong to non-sensitive categories, this approach may provide a better trade-off between user privacy and the Web economy. However, this approach only provides privacy protection against previously defined sensitive content, when tracked through HTML cookies, and thus does not preclude more sophisticated tracking technologies (such as canvas fingerprinting) and less simple tracking methods based on IP address.

8) MyAdChoices: Parra-Arnau et al. [64] propose a web-browser plug-in aimed at bringing transparency over tracking and advertising, and providing a certain level of granularity with regard to blocking ads. As for transparency, the plug-in estimates if the ads delivered to a user may have been generated from their previously visited pages. It also permits users to know if the browsing profiles available to trackers and ad companies may show common or unique interests.

In terms of blocking functionalities, the tool enables users to hide ads by topic category and depending on whether they have been displayed based on users' browsing interests or not. Although the tool provides fine-grained control over ads, it does not prevent any form of tracking; ads are basically hidden to users by applying a black mask on top of ad images. Another limitation of this approach is that the transparency functionalities come at the cost of additional traffic. The reason is due to the fact that, to decide if an ad is profile based, it must revisit the pages browsed by the user in incognito mode.

9) AdReveal: Liu et al. [72] propose an advertising-transparency platform aimed at studying the ads delivered to some artificial profiles, built from the AOL search query data set [73]. The tool is not intended for end-users, unlike MyAdChoices, and provides a framework that aims to study interest-based and contextual advertising at large scale. The platform, which operates offline and is restricted to DoubleClick ads, analyzes two data sets to this end: the interest categories of *all* ads received both in a tracked session and in an incognito-browsing mode. The authors then use a binary classifier to decide if an ad belonging to a certain category is interest-based *or* contextual.

10) XRay: Similarly to AdReveal, XRay [74] propose a transparency platform which tracks the personal data collected by several Web services, and tries to correlate data inputs (e.g., e-mails and search queries) with data outputs (e.g., ads and recommended links). The proposed platform has been tested for the ads displayed on Gmail and relies on the maintenance of a number of *shadow accounts*, that is, replicates of the original account (e.g., an e-mail account), but which differ in a subset of inputs. All these account instances are operated in parallel by the system and are used to compare the outputs received. Intuitively, if an ad is displayed more frequently on those accounts sharing a certain input (e.g., an e-mail), and this ad never shows up in the rest of shadow instances, then this input is likely to be the cause of said ad.

The major limitations of transparency tools such as MyAdChoices, AdReveal and XRay, come from the necessarily simplified model assumed for the ad-delivery process. Evaluating an ad-transparency tool is, besides, extremely challenging since the ground truth of targeting decisions is unknown. XRay, in addition, provide a solution which is not intended for end-users, i.e., it is not designed to be used by a single user who wishes to find out what particular ads are targeted to them.

C. Industrial and Commercial Solutions

Commercial solutions mostly take the form of web-browser extensions. Since all user interactions with the Web are handled through the browser, taking advantage of such an interface to filter or block third-party tracking seems a reasonable approach. These browser extensions endeavor to protect user privacy by blocking third-party interactions. This strategy is

Protection mechanism	Location	Scope	Protection strategy				
			Transparency	Blocking	Obfuscation	Sandboxing	User inclusion
Adblock Plus and similar	local	local	✓	✓			
Ghostery	local	local	✓	✓			
Google Sharing	third-party	local		✓	✓		
Brave	N/A	global (1)	✓	✓			✓
Subscribe2Web (2)	N/A	global		✓(3)			✓
Google Contributor (2)	N/A	global		✓(3)			✓
Network-level ad blocking	local, third-party	local	✓	✓			✓

TABLE VII: Summary of the commercial solutions for privacy protection in online advertising, classified according to the parameters examined in Sec. IV-C. (1) Its global scope feature, which allow users to get involved in the online advertising ecosystem, is still in beta version; (2) devised by Mozilla and Google, these approaches are proposed as future products; (3) as a future project, it proposes reducing/blocking the ads displayed to users; however, it is unknown if the approach taken will imply cutting-off Web tracking or not.

Extension	Blocking strategy	Trust level	Expected performance
Adblock Plus	List-based	Medium	Low
Ghostery	List-based	Low	Medium
AdBlock	List-based	Medium	Low
Disconnect	List-based	High	High
Lightbeam	List-based (items added manually)	High	Medium
Privacy Badger	Heuristic-based/dynamic	High	Medium
DoNotTrackMe/Blur	List-based	Medium	Medium
MyTrackingChoices	Dynamic	High	High
MyAdChoices	Dynamic	High	High
Brave	List-based	High	High

TABLE VIII: Browser-based approaches described in terms of their blocking strategies, trust level and performance.

usually implemented both statically, based on lists of banned trackers, or dynamically, based on heuristics and automatic learning. The specific implemented approach has an immediate, evident effect on the trust level over the tool and even the performance of the browser. For instance, those tools based on large blocking lists (such as the ones available for the most popular browser extensions) may perform worse due to the need to check these lists every time a page is visited. In this regard, we hasten to stress that the criteria employed to manage such lists is not clear at all. This obviously may arise suspicion and reduce the level of trust in these solutions.

A rich variety of browser-based solutions are currently available as commercial products, some of them providing users with control over online advertising. The controversy stirred by the use of the blocking lists they rely on [75] [76], however, has motivated the rise of *open-source, transparency* technologies that may prevent ad companies from interfering. Next, we shall examine a particular class of solutions called ad blockers. Although there exist numerous tools of this kind, our analysis will focus only on the most popular ones, namely Adblock Plus and Ghostery. Other ad blockers such as AbBlock [77], Lightbean [78], Disconnect [79], Blur [80], SuperBlock Adblocker [81], AdRemover [82], AdBlock Pro [83] and uBlock [84] operate similarly.

The last group of (four) initiatives explored in this section are not yet implemented and aim at radically changing the paradigm of the online ad delivery. Sponsored by relevant institutions such as Google, Yahoo and Internet providers, these initiatives propose a user-driven architecture whose main aim is to strike a better trade-off between user privacy and the Web economic model. Table VII shows a classification of the commercial solutions analyzed in the coming subsections on the basis of the protection parameters described in Sec. IV-C. (1). Table VIII shows different aspects of the browser-based proposals (both academic and commercial) such as their strategy to prevent tracking, and the corresponding trust level and performance.

1) Adblock Plus: Available for all major browsers, Adblock Plus is an extension that blocks tracking and ad serving [85] based on filtering lists which specify the elements of a website that may be blocked. These elements include malware domains, banners, pop-up windows, and video ads on Facebook and YouTube. Users enable blocking by adding the filtering lists of their preference, managed in [86]. Adblock Plus is the world's most downloaded ad blocker and therefore the tool that is currently threatening the Internet business model [87].

This ad blocker has recently incorporated a whitelisting mechanism —enabled by default— for nonintrusive ads that meet certain criteria. These criteria are defined in the *acceptable ads* initiative [88], and although the adherence to this initiative is optional for advertisers, much criticism has arisen especially after the revelation that Adblock Plus was getting money from ad companies to whitelist them [89], [90].

2) Ghostery: Developed by Evidon, Ghostery [91] is a proprietary browser add-on capable of detecting third-party trackers. By default, this tool blocks the execution of the tracking cookies as well as the scripts belonging to the tracking companies that are blacklisted. The list in question is elaborated by the company itself. Even though the tracking companies in this list are classified into five categories, according to their different purposes (analytics, web bugs, privacy, advertising, and widgets), it is highly unlikely that users recognize such categories or entities to make a conscious configuration of the tool [92]. However, using such lists may simplify the configuration of the add-on.

When a user browses the Web, Ghostery shows the trackers that are blocked on each page (through a non-intuitive or usable categorization), and offers the possibility of adding any such trackers to a whitelist. Ghostery protects users' privacy from advertisers by blocking scripts, images, objects, and documents embedded by companies the user might not trust. Other tracking mechanisms such as web or canvas fingerprinting are not addressed by Ghostery. Finally, the tool has been criticized for its default behavior [92] which allows Ghostery to collect information about the blocked ads, and afterwards sell it to ad companies [93].

3) Google Sharing: Google Sharing [94] is a system that provides privacy protection by avoiding the tracking conducted by Google. It consists in a Firefox extension that redirects user's requests to an external proxy, where a group of identities associated with cookies are managed. These cookies replace the ones included in original requests, masking a user's identity, and are then forwarded to Google along with the original request. Even when they allow users to send encrypted requests, however, user privacy can still be compromised if collusion exists between the proxy server and Google servers.

4) Brave: It is a web browser —and not a plug-in— that natively embeds functions to block intrusive ads and third-party tracking by default [95]. This proposal allows replacing the ads available on the visited pages with others from Brave's own advertising network, claimed to be less intrusive and more privacy-friendly.

The proposed browser contemplates integrating users into the online advertising business by paying them 15% of the gross ad revenue. In this regard, users are given the option to donate such money to publishers, in exchange for an ad-free browsing experience. Among other transparency functionalities, users may learn the number and type of blocked ads, the trackers present on the visited pages and HTTP redirections.

The upshot is that Brave operates similarly to an entire ad platform, but managed by a single company. The solution completely dispenses with the present advertising infrastructure and aims at building a new one, apparently fairer and more private. However, this approach has sparked much criticism [96] since users' browsing data are collected and processed by a *single* company, which merely shift users' trust from the current multi-system advertising model to this new single entity.

5) Subscribe2Web: Developed by Mozilla, Subscribe2Web [97] endeavors to address some of the privacy risks examined in Sec. III. Based on the idea that online advertising is crucial for the present Web content model, Subscribe2Web looks for a way whereby the main actors (in particular, content creators and users) can meet and have a natural exchange of value. Mozilla's proposal is to eliminate the current Web dependency on ads, in order to fund the content creation by directly compensating content and service providers. The aim is to provide the Web with an API accessible from any browser through which users would pay a monthly subscription in exchange for accessing ad-free content.

6) Google Contributor: Contributor [98] is an initiative supported by Google to reduce the amount of ads delivered by its advertising services. Its main aim is not directly related to protect user privacy but to give users the possibility to eliminate ads from their favorite sites. Because advertisers would be partially excluded by this approach, users registered with this service would have to somewhat support the free ad sites by paying a monthly fee. Thus, Contributor relies on a novel idea where users are considered as active agents in the Web economic model.

7) Network-Level Ad Blocking: Recently, some Internet service providers have started to cooperate with ad companies to implement ad blocking technologies [99] [100]. This is the case of Three, an operator in the UK and Italy, which is working with Shine Technologies to deploy network-based ad blocking.

With these network-level ad blocking practices, a new powerful agent breaks into the online advertising ecosystem, stating that customers should have more control over the content displayed on their browsers, especially when they would be paying for every downloaded byte. Even though not much information is available about the blocking mechanisms to be used, the goal would not be to eliminate advertising but to give users more information (transparency) and the option to decide what to block (control). In the long term, this approach may help protect user privacy, offer relevant and non-intrusive ads, and allow advertisers to take upon the data charges for downloaded ads.

V. Discussion

In Secs. II and III we made it clear that online advertising is a market where the exchanged goods are the users' data. Therefore, the multiple interactions among the entities of such a market might entail privacy risks for its users. Third-party entities from online ad platforms, such as DSPs, SSPs and ad exchanges, and many others offering a transport channel are especially responsible for the collection and aggregation of most of the user information employed as the raw material for their targeted ad delivery strategies.

The main concern of privacy advocates about online advertising is that the user information collected by intermediate entities might be employed to uniquely identify users or classify them in order to, for instance, discriminate their patterns of behavior. This risk is significantly worse due to the following factors specific to the online advertising ecosystem:

- most processes are performed in the background so the infrastructure is not transparent by default for users;
- user data is massively collected by several intermediary entities;
- the user data are necessarily distributed and processed at very high speeds due to the real-time requirements of advertising, which makes it difficult for ad companies to anonymize and protect such data right after their collection;
- cooperation is encouraged between intermediate entities in terms of data sharing;

- multiple items of information can be collected about users through non-consented interactions that are indirectly triggered from users;
- information about users along with processed metadata are commonly exchanged in an unencrypted form between ad serving entities; and
- advanced, resistant and intrusive tracking mechanisms are used to identify users online.

The inevitable consequence of the aforementioned procedures, supporting, in practice, the massive trade of user profiles, is the abusive and nonconsensual identification and classification of users [33] which in extreme cases might entail, for instance, discriminative treatment [58] when they receive online services. These factors of the online advertising ecosystem promote the development of advanced mechanisms to track users through the Web. Practices such as CM, flash cookie setting, canvas fingerprinting, and device fingerprinting in general are massively implemented [30] [47] [48] [44] and sometimes become so intrusive that users are tracked even when some of such fingerprints have been deleted. Most of these practices build on cookies as a mechanism to identify users and to even store information about them. Cookies are commonly combined with other technologies such as canvas and device fingerprinting to obtain a less ephemeral trace of users. Meanwhile, CM exploits the identifying strings retrieved by using cookies to promote massive cooperation among online advertising entities.

The proved complexity of the online advertising ecosystem and the generalized control that huge companies have acquired over ad distribution infrastructures [3] [1] significantly limits the scope of the proposed privacy protection policies. As a consequence, most of the privacy-protecting approaches build on local mechanisms which aim at disabling third-party interactions triggered from the user side to online advertising infrastructures (mainly between users and SSPs), directly blocking user information leakage. Such local approaches are commonly implemented as web browser extensions that provide users with transparency and ad control functionalities [85] [64]. Still located between users and SSPs, other proposals suggest filtering strategies carried out by third-party entities (so-called brokers) [62] [67] [61] which may have access to the interactions directly performed between a group of users and the advertising entities. Given the evident limitations of local approaches, some initiatives have envisioned privacy-by-design advertising platforms where privacy guarantees are provided with a global scope [95] [98] [97]. Interestingly, such initiatives agree on integrating users into the advertising ecosystem.

Our analysis has examined privacy mechanisms with various levels of impact on the Web. To start, offering transparency to users is probably the most appreciated feature of ad blockers (and research platforms such as AdReveal), which is complemented with tracking blocking capabilities to give users a significant level of control. Notwithstanding, the usability of ad blockers for nontechnical users is questionable [92] and these approaches dismiss much, if not all, the current online advertising ecosystem, thus hindering the current economic model of the Web supported by ads.

Even though some of these blocking-based solutions have become pretty popular (e.g., Adblock Plus), the changing business models and default (whitelisting) behaviors of some of these commercial solutions have stirred great controversy. Fortunately, other approaches supported by privacy activists, academics and foundations (such as EFF) are proposing more adequate and usable technologies (e.g., Privacy Badger, MyTrackingChoices) that may block tracking according to users' preferences [63] [64]. Other more refined variants of this blocking strategy are obfuscation and sandboxing [35] [67] [69] [66] (proposed by Obliviad, AdJail, Privad, Adnostic, RePriv and Ghostery). The ultimate aim of these mechanisms is also bounding the amount of user information learned by ad platforms, while striving to adapt to the current advertising business paradigm. As for the privacy threats posed by the structure and capabilities of online advertising, by blocking third-party tracking most commercial solutions claim to hamper cookie setting and thus CM. Canvas fingerprinting can be blocked by most local solutions, yet only on a per-domain basis, the same way as flash cookies. Remarkably, combining at least two ad blockers should offer enough protection against most of the threats described in Sec. III.

Finally, given the dynamic nature of user and ad platform economic incentives [101] [102] with respect to privacy, it seems reasonable to propose new and more private ad distribution (and economic) models. Undoubtedly, this should be with the help of mechanisms that allows users to play a more active role on deciding whether to be tracked or not [95] [97]. Inevitably, this level of control would imply an important reduction in revenue for publishers, and thus require users to directly pay content creators.

Since online privacy may be measured with respect to the interest of Web users to protect their browsing data and that of adversaries to exploit such information, analyzing the respective motivations of the different actors is also of great interest.

Without a doubt, economic incentives have encouraged intermediary entities, advertisers, publishers, and users to participate (consciously or not) in online advertising. Users' unconscious motivation to get involved in online advertising, playing the role of the product, is linked to their need to access free content and services on the Internet. Since the vast majority of Web content and services is paid from advertisement revenue, users have few options to opt-out.

On the other hand, publishers need to help advertisers and ad platforms in their bid to maximize their revenue. For this purpose, website owners are disposed to cede valuable space in their sites and information about their users to such intermediary parties, which thereafter will be responsible for deploying ad-delivery mechanisms. Thus, in exchange for money, publishers surrender some control of its interaction with users and indirectly participate in the disclosure of private contextual information to ad platforms.

On the other side, the interest of advertisers in actively leaking user information is rather reduced, unless several of them collude to share. However, advertisers typically engage ad exchanges and DSPs' services to receive contextual information, which may be useful to deliver targeted ads. Therefore,

	Threats							User role			
	T1. First-party tracking (1)	**T2**. Third-party tracking	**T3**. Cookie matching	**T4**. Fingerprinting	**T5**. Flash cookies	**T6**. Canvas fingerprinting	**T7**. HTML5 local storage	**L1**. Lack of awareness	**L2**. Lack of control	**L3**. Bounded knowledge	**Observations**
ObliviAd		✓									Obfuscating user preferences may prevent third-party tracking. But IP address, user agents and other content embedded in websites might still be used as sources of fingerprinting
AdJail		N/A	N/A	N/A	N/A	N/A	N/A				Provides security services such as integrity and confidentiality in the publisher side
Privad		✓									May avoid third-party tracking, but other user data such IP address and certain user agent may be used as sources of fingerprinting
Adnostic		✓							✓	✓	If enforced by ad platforms, it would discourage third-party tracking. Protection against other threats is not considered
RePriv		N/A	N/A	N/A	N/A	N/A	N/A		✓	✓	Users may control their browsing data on their side, but nothing may prevent external tracking
Privacy Badger		✓	✓		✓	✓	✓	✓		✓	Blocks most tracking mechanisms, but little control is given to users
MyTrackingChoices		✓			✓			✓	✓	✓	Users may block third-party tracking on a more granular level, but protection is against previously defined sensitive content
MyAdChoices								✓	✓	✓	Does not prevent any form of tracking, and ads are hidden from the user, not blocked
AdReveal								✓			Framework aimed at studying interest-based and contextual advertising at large scale
XRay								✓			Platform, not intended for end users

TABLE IX: Online advertising privacy threats and academic proposals addressing them. (1) Since this threat derives directly from interactions explicitly triggered by users, protecting from it is a challenging task.

advertisers' incentives to collect user information are high as well.

The commercial nature of online advertising has spurred a debate about the motivation of the involved entities to protect privacy and to profit from user data. Although apparently opposed, the motivations of users and advertising intermediaries for privacy might vary according to factors that are not commonly considered. Research on the economic behavior of data holders in the market of online advertising [103] has shown that an increased level of user-targeting can reduce their profit due to an exacerbated transfer of value to advertisers. Specifically, advertisers would be gradually less interested in bidding for user impressions as more detailed information is given to them. That way, according to Bergemann and Bonatti [102], an unexpected incentive may appear for data holders to provide reduced accuracy in the exchanged user data, with the aim of generating greater demand from advertisers and thus greater profit for data holders. Interestingly, such increase in profit may lead to more privacy for users (given by the reduced precision of user data leaked to advertisers). Nonetheless, a recent study by Taylor and Wagman [104] poses that the effects of targeting capabilities on profits depend on market and is, consequently, given by context.

Users seem to face a similar contextual dichotomy even though the concern about privacy is generalized [105] [106]. The fact is that the creation of a marketplace in personal data may shift the balance of power between individuals and companies that gather data. According to some recent studies, this is a shift people would be willing to embrace. Just over half of the 9 000 people surveyed worldwide said they would share data about themselves with companies in exchange for cash [107]. A separate survey has found that 42 percent of more than a thousand 13-17-year-olds in the U.K would rather accept cash for their personal data than earn money from a job [108]. Lastly, it was reported in [109] that 56 percent of the consumers surveyed would be willing to give up personal data provided that they received some kind of economic compensation. This dichotomy between users' concerns and intentions regarding privacy might obey, according to Acquisiti et al. [101], to multidimensional factors relative to the context where the user operates, such as their lack of awareness about privacy risks, and cognitive and behavioral biases. The upshot is that users' assessment of their own privacy will strictly shape the impact of external threats. Tables IX and X summarize our discussion of privacy technologies and how they may address the threats identified in Sec. III.

A. Future Research Directions

The complexity of online advertising poses various challenges to user privacy. From the analysis conducted in the previous

	Threats							User role			
	T1. First-party tracking (1)	**T2**. Third-party tracking	**T3**. Cookie matching	**T4**. Fingerprinting	**T5**. Flash cookies	**T6**. Canvas fingerprinting	**T7**. HTML5 local storage	**L1**. Lack of awareness (2)	**L2**. Lack of control (3)	**L3**. Bounded knowledge (4)	**Observations**
AdBlock Plus		✓	✓		✓			✓	✓	✓	Protects against some of the analyzed privacy threats, but threatens the economic model of the Web
Ghostery		✓	✓		✓			✓	✓	✓	Offers additional transparency functionalities to users regarding third-party tracking
Google Sharing		✓	✓								Aimed at protecting users only from cookie tracking performed by Google
Brave		✓		✓(5)	✓	✓		✓	✓	✓	Based on the paradigm of a more user-driven ad platform; offers transparency and a great level of control to users
Subscribe2Web		N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Posits a future ad-free Web, but it is unknown if this commercial solution necessarily implies stopping tracking users
Google Contributor		N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Proposes an alternative economic model for the Web, but it is not clear if users will support content creators economically
Network-level ad blocking		✓	✓		✓	✓	✓	✓	✓		At network level, Internet providers are capable of offering similar blocking services than those of ad blockers

TABLE X: Online advertising privacy threats and commercial solutions addressing them. (1) Since this threat derives directly from interactions explicitly triggered by users, protecting from it is a challenging task; (2) most solutions only show the number of trackers detected/blocked; (3) control is commonly enforced by given users blocking capabilities; (4) default configurations simplify their use but at the expense of privacy; (5) fingerprinting protection is currently in beta version.

sections, we envisage two main lines of future research: identifying new privacy threats and providing new protection mechanisms. Given the opacity of ad platforms, we believe that further exploring the tracking capabilities of the advertising industry will help discover their potential to become privacy attackers. But not only that, unveiling the user data exchange processes within ad platforms would expose the extent to which some of their intermediate entities are prone to become massive surveillance agents. A better knowledge of the adversary will contribute to develop protection mechanisms which are more tailored to the above mentioned privacy threats.

As for privacy mechanisms, a natural next step would be combining some of the proposals described in Sec. IV. Such synergy shall generate more robust and useful privacy solutions for detecting user-related flaws and invasive tracking behaviors, and better adapting privacy enhancing technologies to the current Web economic model.

Regarding the strategy posed by current privacy protection approaches (namely blocking, obfuscation, sandboxing, and user inclusion), a further analysis on their impact on the Web economic model will reveal if such proposals are effectively adapting to the current advertising business model, without a significant side effect.

A further research direction for improving users' privacy in online advertising is to create smarter protection tools in the user side, that is, developed as browser complements. Intelligibility, usability and flexibility are some of the parameters that need to be considered to enable mechanisms to give users real transparency and control over their browsing data. In this regard, a great deal of work has to be done to develop tools that let users effectively enforce their motivations on the protection strategy selected.

Another strand of research may consider the scope of the protection strategy, currently limited to the user side. Extending the scope of the privacy protection mechanisms to the different players (e.g., publishers, advertisers, ad exchanges) might result in a more solid approach. Accordingly, analyzing and evaluating the privacy policies and protection mechanisms offered by ad platforms might contribute to detect their flaws and make improvements.

To go beyond the simplistic (and endangering) blocking strategy of some approaches examined in Sec. IV, new advertising models have to be envisioned that provide flexible two-way communication interfaces between users and ad platforms through which they could directly manage their relationship according to their interests. While economic interests of advertising entities are widely known, user motivations related to privacy, advertising choices and even economic incentives should be seriously considered by such models. Undoubtedly, more transparent and balanced interactions will derive in an increased sense of security and thus of privacy.

A more user-driven advertising platform, where user interests regarding their privacy and profit may be variable (not always opposing to the advertisers'), and the assessment of user information as an asset with intrinsic economic value, not only for intermediate advertising entities, but also for users, will help to study the trade-off between such value and the privacy of users involved in online advertising transactions.

VI. Conclusions

Online advertising has become ubiquitous on the Internet and the revenues ad serving generates for publishers are supporting the existing free Internet access model. As a consequence of such ubiquity, online advertising has triggered the creation of a massive transport channel whose intermediary components have access to billions of users and, in particular, to their data. Even though gigabytes of aggregated user data support more targeted advertising campaigns, the inherent lack of transparency of online advertising entails serious risks to user privacy. In this article, by breaking down the instances of online advertising platforms and their corresponding capabilities (regarding user data), we have outlined an attacker model to describe the potential hazards to user privacy. We have emphasized the variety of information subject to be collected, the large number of intermediaries involved, their advanced and intrusive tracking capabilities, and the impact of advertising practices on privacy.

Unlike what happens with other online privacy threats, there is little users can do to completely prevent risks coupled with online advertising. Nevertheless, several solutions are offered to help protect the privacy of users within such an opaque ecosystem. Accordingly, we have offered a wide range of mechanisms in this paper and we classified them into local solutions (browser and third-party based) and proposals based on new ad serving paradigms. On the one hand, some of the local solutions are very popular and their blocking approaches are already negatively impacting the economic model of online advertising. On the other hand, new advertising models are arising to offer native privacy and a stronger role for the user, while still proposing radical variations of the current advertising logic.

In addition, we have elaborated on the pros and cons of some of the aforementioned protection mechanisms with regard to the threats they try to alleviate within online advertising platforms. In such analysis, we also outlined dynamic and smarter approaches proposed to avoid radical blocking mechanisms. Yet, based on the proposals analyzed, we have found it very hard to provide more privacy in the online advertising ecosystem without significantly modifying the ad delivery model to give users more control and to reduce the financial dependence of Internet content on advertising.

Acknowledgment

We would also like to thank the anonymous reviewers for their immensely helpful suggestions to improve the readability and contents of this paper. This work was partly funded by the European Commission through the project H2020-644024 "CLARUS", the Spanish Ministry of Economy, Industry and Competitiveness (MINECO) through the projects TEC2014-54335-C4-1-R "INRISCO", TIN2016-80250-R "Sec-MCloud", as well as by the Government of Catalonia under grants 2014 SGR 1504 and 2014 SGR 00537. J. Parra-Arnau is the recipient of a Juan de la Cierva postdoctoral fellowship, FJCI-2014-19703, from the MINECO. José Estrada-Jiménez and Ana Rodríguez-Hoyos acknowledge the support from Escuela Politécnica Nacional (EPN) for their doctoral studies at Universitat Politècnica de Catalunya (UPC).

References

[1] S. Yuan, J. Wang, and X. Zhao, "Real-time bidding for online advertising: measurement and analysis," in *Proceedings of the Seventh International Workshop on Data Mining for Online Advertising*. ACM, 2013, p. 3.

[2] H. Beales, "The value of behavioral targeting," Netw. Advertising Initiative, Tech. Rep., Mar. 2010, accessed on 2016-01-15. [Online]. Available: http://www.networkadvertising.org/pdfs/Beales_NAI_Study.pdf

[3] M. Smith, *Targeted: How technology is revolutionizing advertising and the way companies reach consumers*. AMACOM Div American Mgmt Assn, 2014.

[4] "Real-time bidding protocol - processing the request," accessed on 2016-04-07. [Online]. Available: https://developers.google.com/ad-exchange/rtb/request-guide

[5] S. Yuan, A. Z. Abidin, M. Sloan, and J. Wang, "Internet advertising: An interplay among advertisers, online publishers, ad exchanges and web users," *arXiv: 1206.1754*, 2012, arXiv preprint.

[6] A. M. McDonald and L. F. Cranor, "Americans' attitudes about internet behavioral advertising practices," in *Proceedings of the 9th annual ACM workshop on Privacy in the electronic society*. ACM, 2010, pp. 63–72.

[7] D. Lyon, Ed., *Surveillance as Social Sorting: Privacy, Risk and Automated Discrimination*. Routledge, Dec. 2002.

[8] U. Senate, "Online advertising and hidden hazards to consumer security," U.S. Senate, Tech. Rep., 2014. [Online]. Available: http://www.hsgac.senate.gov/hearings/online-advertising-and-hidden-hazards-to-consumer-security-and-data-privacy

[9] D. S. Evans, "The online advertising industry: Economics, evolution, and privacy," *The journal of economic perspectives*, vol. 23, no. 3, pp. 37–60, 2009.

[10] "The cost of ad blocking," PageFair, Res. Rep., Aug. 2015.

[11] "The state of online advertising," Adobe, Tech. Rep., 2012, accessed on 2015-09-11. [Online]. Available: http://www.adobe.com/aboutadobe/pressroom/pdfs/Adobe_State_of_Online_Advertising_Study.pdf

[12] G. Marvin, "Consumers now notice retargeted ads," Marketing Land, Tech. Rep., Dec. 2013, accessed on 2015-08-12. [Online]. Available: http://marketingland.com/3-out-4-consumers-notice-retargeted-ads-67813

[13] F. Rejón-Guardia and F. J. Martínez-López, "Online advertising intrusiveness and consumers avoidance behaviors," in *Handbook of Strategic e-Business Management*. Springer, 2014, pp. 565–586.

[14] M. Schudson, *Advertising, the uneasy persuasion (RLE Advertising): Its dubious impact on American society*. Routledge, 2013.

[15] D. Morgan, "TV audience fragmentation is an inescapable reality: Embrace it," Sep. 2012, accessed on 2016-03-04. [Online]. Available: http://www.mediapost.com/publications/article/182946/tv-audience-fragmentation-is-an-inescapable-realit.html

[16] K. Nelson-Field and E. Riebe, "The impact of media fragmentation on audience targeting: An empirical generalisation approach," *Journal of Marketing Communications*, vol. 17, no. 01, pp. 51–67, 2011.

[17] K. McSpadden, "You now have a shorter attention span than a goldfish," May 2015, accessed on 2016-03-04. [Online]. Available: http://time.com/3858309/attention-spans-goldfish/

[18] P. Minnium, "8 reasons why digital advertising works for brands," Nov. 2014, accessed on 2016-03-12. [Online]. Available: http://marketingland.com/10-reasons-digital-advertising-works-brands-108151

[19] A. A. Kardan and M. Hooman, "Targeted advertisement in social networks using recommender systems," in *e-Commerce in Developing Countries: With Focus on e-Security (ECDC), 2013 7th Intenational Conference on*. IEEE, 2013, pp. 1–13.

[20] J. Yan, N. Liu, G. Wang, W. Zhang, Y. Jiang, and Z. Chen, “How much can behavioral targeting help online advertising?” in *Proceedings of the 18th international conference on World wide web*. ACM, 2009, pp. 261–270.

[21] C. E. Tucker, “Social networks, personalized advertising, and privacy controls,” *Journal of Marketing Research*, vol. 51, no. 5, pp. 546–562, 2014.

[22] C. Gayomali, “It would cost each user $232 a year for an ad-free internet, study finds,” Aug. 2014, accessed on 2016-02-27. [Online]. Available: http://www.fastcompany.com/3034670/fast-feed/it-would-cost-each-user-232-a-year-for-an-ad-free-internet-study-finds

[23] IAB, “Digital ad revenues surge 19 percent, climbing to $27.5 billion in first half of 2015,” Oct. 2015, accessed on 2016-03-06. [Online]. Available: http://www.iab.com/news/digital-ad-revenues-surge-19-climbing-to-27-5-billion-in-first-half-of-2015-according-to-iab-internet-advertising-revenue-report/

[24] OpenX, “Ad networks vs. ad exchanges: How they stack up,” Jul. 2010, accessed on 2016-03-06. [Online]. Available: http://openx.com/blog/openx-releases-new-whitepaper-ad-networks-vs-ad-exchanges/

[25] J. R. Mayer and J. C. Mitchell, “Third-party web tracking: Policy and technology,” in *2012 IEEE Symposium on Security and Privacy*. IEEE, 2012, pp. 413–427.

[26] Google, “Cookie matching,” Mar. 2016, accessed on 2016-03-07. [Online]. Available: https://developers.google.com/ad-exchange/rtb/cookie-guide#examples

[27] D. Laffey, “Paid search: The innovation that changed the web,” *Business Horizons*, vol. 50, no. 3, pp. 211–218, 2007.

[28] S. Horowitz, “Real time bidding - rtb - programmatic advertising explained,” accessed on 2016-03-09. [Online]. Available: http://executive-digital.com/digital-marketing-experts/what-is-real-time-bidding-and-why-is-it-more-effective-than-direct-bidding-methods/

[29] L. Olejnik, T. Minh-Dung, and C. Castelluccia, “Selling off privacy at auction,” in *Proc. IEEE Symp. Netw. Distrib. Syst. Secur. (SNDSS)*, Feb. 2014.

[30] A. Ghosh, M. Mahdian, R. P. McAfee, and S. Vassilvitskii, “To match or not to match: Economics of cookie matching in online advertising,” *ACM Transactions on Economics and Computation*, vol. 3, no. 2, p. 12, 2015.

[31] A. Goldfarb and C. E. Tucker, “Privacy regulation and online advertising,” *Management science*, vol. 57, no. 1, pp. 57–71, 2011.

[32] B. Ur, P. G. Leon, L. F. Cranor, R. Shay, and Y. Wang, “Smart, useful, scary, creepy: perceptions of online behavioral advertising,” in *proceedings of the eighth symposium on usable privacy and security*. ACM, 2012, p. 4.

[33] J. Parra-Arnau, D. Rebollo-Monedero, and J. Forné, “Measuring the privacy of user profiles in personalized information systems,” *Future Generation Computer Systems*, vol. 33, pp. 53–63, 2014.

[34] S. Pandey, M. Aly, A. Bagherjeiran, A. Hatch, P. Ciccolo, A. Ratnaparkhi, and M. Zinkevich, “Learning to target: What works for behavioral targeting,” in *Proc. Int. Conf. Inform., Knowl. Manage. (CIKM)*. ACM, 2011, pp. 1805–1814.

[35] V. Toubiana, A. Narayanan, D. Boneh, H. Nissenbaum, and S. Barocas, “Adnostic: Privacy preserving targeted advertising,” in *Proc. IEEE Symp. Netw. Distrib. Syst. Secur. (SNDSS)*, Feb. 2010, pp. 1–21.

[36] P. Eckersley, “How unique is your web browser?” in *International Symposium on Privacy Enhancing Technologies Symposium*. Springer, 2010, pp. 1–18.

[37] “How real a threat is de-anonymization?” May 2011, accessed on 2016-03-09. [Online]. Available: https://swildstrom.wordpress.com/2011/05/31/how-real-a-threat-is-de-anonymization/

[38] A. Narayanan and V. Shmatikov, “Robust de-anonymization of large sparse datasets,” in *Proc. IEEE Symp. Secur., Priv. (SP)*. IEEE Comput. Soc., May 2008, pp. 111–125.

[39] C. A. Dwyer, “Behavioral targeting: A case study of consumer tracking on levis. com,” *Available at SSRN 1508496*, 2009.

[40] M. A. Bashir, S. Arshad, W. Robertson, and C. Wilson, “Tracing information flows between ad exchanges using retargeted ads,” in *Proceedings of the 25th USENIX Security Symposium*, 2016.

[41] F. Roesner, C. Rovillos, T. Kohno, and D. Wetherall, “Balancing privacy and functionality of third-party social widgets,” *USENIX Magazine*, 2012.

[42] N. Vratonjic, M. H. Manshaei, J. Grossklags, and J.-P. Hubaux, “Adblocking games: Monetizing online content under the threat of ad avoidance,” in *The Economics of Information Security and Privacy*. Springer, 2013, pp. 49–73.

[43] L. Olejnik and C. Castelluccia, “To bid or not to bid? measuring the value of privacy in RTB,” accessed on 2016-05-21. [Online]. Available: http://lukaszolejnik.com/rtb2.pdf

[44] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz., “The web never forgets: Persistent tracking mechanisms in the wild,” in *Proc. ACM Conf. Comput., Commun. Secur. (CCS)*, Washington, DC, Nov. 2014, pp. 674–689.

[45] S. Englehardt, D. Reisman, C. Eubank, P. Zimmerman, J. Mayer, A. Narayanan, and E. W. Felten, “Cookies that give you away: The surveillance implications of web tracking,” in *Proceedings of the 24th International Conference on World Wide Web*. ACM, 2015, pp. 289–299.

[46] S. Englehardt and A. Narayanan, “Online tracking: A 1-million-site measurement and analysis,” in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2016, pp. 1388–1401.

[47] A. Soltani, S. Canty, Q. Mayo, L. Thomas, and C. J. Hoofnagle, “Flash cookies and privacy,” in *Proc. AAAI Spring Symp. Intell. Inform. Priv. Manage.* Assoc. Adv. Artif. Intell., 2010.

[48] M. D. Ayenson, D. J. Wambach, A. Soltani, N. Good, and C. J. Hoofnagle, “Flash cookies and privacy II: Now with HTML5 and ETag respawning,” *Available at SSRN 1898390*, 2011.

[49] K. O. Advertising, “Flash cookies, HTML 5 storage and HTTP cookies - know online advertising,” accessed on 2016-06-15. [Online]. Available: http://www.knowonlineadvertising.com/difference-between/flash-cookies-html-5-storage-and-http-cookies

[50] A. M. McDonald and L. F. Cranor, “Survey of the use of adobe flash local shared objects to respawn HTTP cookies,” *ISJLP*, vol. 7, p. 639, 2011.

[51] Addthis.com, “Get more like, shares and follows with smart website tools,” Mar. 2016, accessed on 2016-05-30. [Online]. Available: http://www.addthis.com

[52] C. J. Hoofnagle, A. Soltani, N. Good, D. J. Wambach, and M. D. Ayenson, “Behavioral advertising: the offer you cannot refuse,” 2012.

[53] D. Kravets, “Lawsuit targets mobile advertiser over sneaky html5 pseudo-cookies,” Sep. 2010, accessed on 2016-05-30. [Online]. Available: https://www.wired.com/2010/09/html5-safari-exploit

[54] M. S. Ackerman, L. F. Cranor, and J. Reagle, “Privacy in e-commerce: examining user scenarios and privacy preferences,” in *Proceedings of the 1st ACM conference on Electronic commerce*. ACM, 1999, pp. 1–8.

[55] L. Agarwal, N. Shrivastava, S. Jaiswal, and S. Panjwani, “Do not embarrass: re-examining user concerns for online tracking and advertising,” in *Proceedings of the Ninth Symposium on Usable Privacy and Security*. ACM, 2013, p. 8.

[56] J. P. Carrascal, C. Riederer, V. Erramilli, M. Cherubini, and R. de Oliveira, “Your browsing behavior for a big mac: Economics of personal information online,” in *Proceedings of the 22nd international conference on World Wide Web*. ACM, 2013, pp. 189–200.

[57] A. E. Schlosser, S. Shavitt, and A. Kanfer, “Survey of internet users? attitudes toward internet advertising,” *Journal of interactive marketing*, vol. 13, no. 3, pp. 34–54, 1999.

[58] L. Sweeney, “Discrimination in online ad delivery,” *Queue*, vol. 11, no. 3, p. 10, 2013.

[59] M. Smith, *Targeted: How Technology Is Revolutionizing Advertising and the Way Companies Reach Consumers*, 1st ed. New York: AMACOM, Nov. 2014.

[60] “US programmatic ad spend tops $10 billion this year, to double by 2016,” eMarketer, Tech. Rep., Oct. 2014. [Online]. Available: http://www.emarketer.com/Article/US-Programmatic-Ad-Spend-Tops-10-Billion-This-Year-Double-by-2016/1011312

[61] Privoxy.org, “Privoxy,” Mar. 2016, accessed on 2016-03-17. [Online]. Available: http://www.privoxy.org

[62] M. Backes, A. Kate, M. Maffei, and K. Pecina, “Obliviad: Provably secure and practical online behavioral advertising,” in *2012 IEEE Symposium on Security and Privacy*. IEEE, 2012, pp. 257–271.

[63] Electronic Frontier Foundation, “Privacy badger,” Mar. 2016, accessed on 2016-03-15. [Online]. Available: https://www.eff.org/privacybadger

[64] J. Parra-Arnau, J. P. Achara, and C. Castelluccia, “MyAdChoices: Bringing transparency and control to online advertising,” *ACM Trans. Web*, 2016, to appear. [Online]. Available: https://hal.inria.fr/hal-01270186/document

[65] D. C. Howe and H. Nissenbaum, *Lessons from the Identity Trail: Privacy, Anonymity and Identity in a Networked Society*. NY: Oxford Univ. Press, 2009, ch. TrackMeNot: Resisting surveillance in Web search, pp. 417–436. [Online]. Available: http://mrl.nyu.edu/~dhowe/trackmenot

[66] M. Fredrikson and B. Livshits, "Repriv: Re-envisioning in-browser privacy," in *Proc. IEEE Symp. Security, Privacy (SP)(May 2011)*, 2010.

[67] S. Guha, B. Cheng, and P. Francis, "Privad: practical privacy in online advertising," in *USENIX conference on Networked systems design and implementation*, 2011, pp. 169–182.

[68] J. P. Achara, J. Parra-Arnau, and C. Castelluccia, "MyTrackingChoices: Pacifying the ad-block war by enforcing user privacy preferences," in *Proc. Annual Workshop Econ. Inform. Secur. (WEIS)*, Jul. 2016.

[69] M. Ter Louw, K. T. Ganesh, and V. Venkatakrishnan, "Adjail: Practical enforcement of confidentiality and integrity policies on web advertisements." in *USENIX Security Symposium*, 2010, pp. 371–388.

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

[71] 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*, ser. NSDI'12. Berkeley, CA, USA: USENIX Association, 2012, pp. 12–12. [Online]. Available: http://dl.acm.org/citation.cfm?id=2228298.2228315

[72] B. Liu, A. Sheth, U. Weinsberg, J. Chandrashekar, and R. Govindan, "Adreveal: Improving transparency into online targeted advertising," in *Proc. Hot Topics in Netw.* ACM, 2013, pp. 12:1–12:7.

[73] "AOL search data scandal," Aug. 2006, accessed on 2013-11-15. [Online]. Available: http://en.wikipedia.org/wiki/AOL_search_data_scandal

[74] M. Lecuyer, G. Ducoffe, F. Lan, A. Papancea, T. Petsios, R. Spahn, A. Chaintreau, and R. Geambasu, "XRay: Enhancing the web's transparency with differential correlation," in *Proc. Conf. USENIX Secur. Symp.*, Aug. 2014.

[75] Ghostery, "How does ghostery make money from the browser extension?" accessed on 2015-12-21. [Online]. Available: https://www.ghostery.com/support/faq/ghostery-add-on/how-does-ghostery-make-money-from-the-add-on/

[76] BetaNews, "Adblock plus updates acceptable ads and reveals how it makes money," 2015, accessed on 2015-12-21. [Online]. Available: http://betanews.com/2015/12/16/adblock-plus-updates-acceptable-ads-and-reveals-how-it-makes-money/

[77] "AdBlock," Mar. 2016, accessed on 2016-03-13. [Online]. Available: https://getadblock.com/

[78] Mozilla, "Lightbeam for firefox," accessed on 2015-12-16. [Online]. Available: https://www.mozilla.org/en-US/lightbeam/

[79] Disconnect.me, "Private browsing," accessed on 2015-12-16. [Online]. Available: https://disconnect.me/disconnect

[80] Abine, "Blur: Keep your web activity and personal info private," Mar. 2016, accessed on 2016-03-15. [Online]. Available: https://dnt.abine.com/#dashboard

[81] "SuperBlock Adblocker," accessed on 2016-05-12. [Online]. Available: https://chrome.google.com/webstore/detail/superblock-adblocker/miijbmhjndcihicbljlcieiajhemmdeb

[82] "AdRemover," accessed on 2016-05-11. [Online]. Available: https://chrome.google.com/webstore/detail/adremover-for-google-chro/mcefmojpghnaceadnghednjhbmphipkb

[83] "Adblock Pro," accessed on 2016-05-11. [Online]. Available: https://chrome.google.com/webstore/detail/adblock-pro/ocifcklkibdehekfnmflempfgjhbedch

[84] "uBlock," accessed on 2016-05-11. [Online]. Available: https://www.ublock.org/

[85] "Adblock Plus - surf the Web without annoying ads!" Nov. 2015, accessed on 2015-11-15. [Online]. Available: https://adblockplus.org

[86] "Easylist - overview," Mar. 2016, accessed on 2016-05-30. [Online]. Available: https://easylist.github.io

[87] PageFair, "The 2015 ad blocking report," 2015, accessed on 2015-11-20. [Online]. Available: https://blog.pagefair.com/2015/ad-blocking-report/

[88] "Allowing acceptable ads in adblock plus," Mar. 2016, accessed on 2016-03-13. [Online]. Available: https://adblockplus.org/en/acceptable-ads

[89] R. Cookson, "Google, Microsoft and Amazon pay to get around ad blocking tool," Feb. 2015, accessed on 2014-03-10. [Online]. Available: http://www.ft.com/cms/s/0/80a8ce54-a61d-11e4-9bd3-00144feab7de.html

[90] M. Sullivan, "Adblock plus was in NYC last week waving the olive branch at advertisers," Nov. 2015, accessed on 2015-11-20. [Online]. Available: http://venturebeat.com/2015/11/13/adblock-plus-was-in-nyc-last-week-waving-the-olive-branch-at-advertisers

[91] "Ghostery - take control of your digital experience," Nov. 2015, accessed on 2015-11-15. [Online]. Available: https://www.ghostery.com/

[92] P. Leon, B. Ur, R. Shay, Y. Wang, R. Balebako, and L. Cranor, "Why Johnny can't opt out: a usability evaluation of tools to limit online behavioral advertising," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 589–598.

[93] A. Henry, "Ad-blocker ghostery actually helps advertisers, if you."

[94] "GoogleSharing." [Online]. Available: www.googlesharing.net

[95] Brave, "Brave software," Mar. 2016, accessed on 2016-05-30. [Online]. Available: https://www.brave.com

[96] T. Adams, "Can the Web save the press from oblivion?" Apr. 2016, accessed on 2016-05-30. [Online]. Available: http://www.theguardian.com/media/2016/apr/17/can-internet-save-printed-press-blendle-lumi

[97] Mozilla, "Subscribe2web," accessed on 2015-11-21. [Online]. Available: https://air.mozilla.org/subscribe2web/

[98] Google, "Google contributor," Mar. 2016, accessed on 2016-03-15. [Online]. Available: https://www.google.com/contributor/welcome

[99] N. Lomas, "Shine signs first european carriers to its network-level ad blocking tech," Feb. 2016, accessed on 2016-03-17. [Online]. Available: http://techcrunch.com/2016/02/18/shine-bags-first-european-carrier-as-three-uk-deploys-network-level-ad-blocking

[100] ——, "U.K. carrier EE looking at giving users control over mobile ads," Nov. 2015, accessed on 2016-03-17. [Online]. Available: http://techcrunch.com/2015/11/23/u-k-carrier-ee-looking-at-giving-users-control-over-mobile-ads

[101] A. Acquisti, C. R. Taylor, and L. Wagman, "The economics of privacy," *Available at SSRN 2580411*, 2016.

[102] D. Bergemann and A. Bonatti, "Selling cookies," *American Economic Journal: Microeconomics*, vol. 7, no. 3, pp. 259–294, 2015.

[103] A. De Corniere, "Search advertising," *Available at SSRN 1967102*, 2013.

[104] C. Taylor and L. Wagman, "Consumer privacy in oligopolistic markets: Winners, losers, and welfare," *International Journal of Industrial Organization*, vol. 34, pp. 80–84, 2014.

[105] J. Turow, J. King, C. J. Hoofnagle, A. Bleakley, and M. Hennessy, "Americans reject tailored advertising and three activities that enable it," *Available at SSRN 1478214*, 2009.

[106] L. Rainie, S. Kiesler, R. Kang, M. Madden, M. Duggan, S. Brown, and L. Dabbish, "Anonymity, privacy, and security online," *Pew Research Center*, vol. 5, 2013.

[107] A. W. Sile, "Privacy compromised? might as well monetize," Jan. 2015.

[108] "The age of digital enlightenment," Logicalis, Tech. Rep., Mar. 2016, accessed on 2016-05-17. [Online]. Available: http://www.uk.logicalis.com/globalassets/united-kingdom/microsites/real-time-generation/realtime-generation-2016-report.pdf

[109] "Privacy and security in a connected life: A study of us, european and japanese consumers," Ponemon Institute, Tech. Rep., Mar. 2015, accessed on 2016-05-14. [Online]. Available: http://www.trendmicro.com/vinfo/us/security/news/internet-of-things/internet-of-things-connected-life-security