

Measuring Online Tracking and Privacy Risks on Ecuadorian Websites

José Estrada-Jiménez^{*†}, Ana Rodríguez-Hoyos^{*†}, Javier Parra-Arnau[‡] and Jordi Forné[†]

^{*} Departamento de Electrónica, Telecomunicaciones y Redes de Información, Escuela Politécnica Nacional (EPN), Ladrón de Guevara, E11-253 Quito, Ecuador.

[†] Department of Telematics Engineering, Universitat Politècnica de Catalunya (UPC), C. Jordi Girona 1-3, E-08034 Barcelona, Spain.

[‡] Department of Computer Engineering and Mathematics, Universitat Rovira i Virgili, Av. Països Catalans 26, E-43007 Tarragona, Catalonia.

Email: ^{*}jose.estrada@epn.edu.ec, ^{*}ana.rodriguez@epn.edu.ec, [‡]javier.parra@urv.cat, [†]jforne@entel.upc.edu

Abstract—Online tracking has become a great enabler of massive surveillance so it is now a critical vector for threatening the privacy of users. Despite the benefits of online tracking for personalized advertising, the complexity of the involved platforms makes it a threat for democracy. In this work, online tracking is measured in Ecuador, a country with a developing adoption of online advertising technologies, having the highest Internet penetration rate in Latin America, but lacking regulation for privacy. By finding out the third party connections triggered through the most popular Ecuadorian websites, the concentration of online tracking is measured in Ecuador. Its impact is also analyzed by studying some particularities in government websites, the usage of advanced mechanisms of tracking, and the adoption of transparency practices in advertising platforms. Our final aim is exposing potential privacy violations.

Index Terms—online tracking, digital advertising, privacy, Ecuador, transparency

I. INTRODUCTION

Browsing the Internet becomes more and more a daily activity such as walking in the street. When walking alone, people have a natural expectancy of privacy that would make them refuse, e.g., any obvious attempt to follow them closely. However, while the same discomfort may arise when “walking” on the Web, there is very little evidence available for users to realize the latent pervasiveness of online tracking. In fact, the opacity and complexity of the Web hide a myriad of interactions triggered by a single user HTTP request. Moreover, many of such interactions among different actors disclose granular user information to third parties, which brings serious concerns regarding the privacy of billions of users. Namely, the “trail” left when browsing a web page is not only known by the visited site, but it is also collected by other entities that “follow” or track users wherever they browse. Said tracking enables third parties to collect a bunch of user data, which is used to build profiles that fuel powerful information systems such as online advertising platforms.

Personalized online advertising is responsible for much of the online tracking performed over users. Online advertising platforms are supported by sophisticated personalization systems that tailor ad content according to the preferences of users; these preferences are learned from the information

collected by tracking. In this line, the more information is collected, the better the performance of personalization systems, and the higher the profits of the advertising platforms. Since online advertising has become a millionaire business [4] that, apparently supports the very existence of the Internet [5], there is a great motivation from multiple instances to collect more and more data, which implies massifying and improving online tracking.

With the involvement of so much user information, online tracking (particularly the one promoted by online advertising) raises serious privacy concerns. The information collected may be so varied and detailed (e.g., location, interests, voting preferences) and the technology used so specialized that tracking may enable third parties to characterize a significant part of a user’s life. Furthermore, state data is currently being collected, due to a real-time mechanism that binds online tracking with every single user web request, enabling third parties (not only Internet providers) to literally monitor each of the user “movements” on the Web.

Due to its prevalence on the Web, measuring online tracking is a great way to characterize the privacy risks of Internet users. The severity of said risks may be illustrated through different indicators such as the level of exposition of user interactions to third parties, the concentration of user information on a few advertising companies, the dynamic behavior of tracking for websites belonging to certain categories, or the suspicious requests to third parties triggered when accessing government web sites.

While related work [9] has performed more general approaches by studying online tracking through the most popular sites of the Web, our analysis focuses on a more reduced scenario represented by the top sites in Ecuador. Ecuador is a small South American country where the right to privacy is guaranteed by its Constitution [6] but where no regulation is applied. Ironically, radical transparency regulation [8] is actively enforced to such extent that the privacy of public employees is put at risk [7]. These dichotomies arise in developing countries whose perceptions regarding privacy are not mature, in part, because of the still incipient adoption of modern personalization systems (including online advertising).

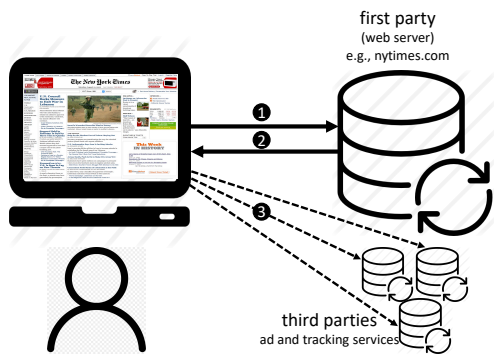


Fig. 1. Requests to third parties (3) triggered by a single HTTP user request (1). When a user browses a website, a redirection command is commonly sent in the HTTP response (2) to spawn further connections to third parties.

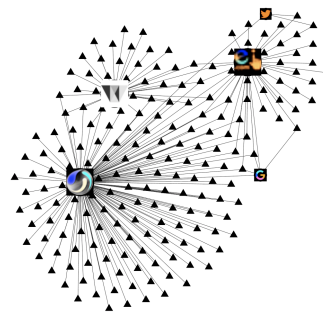


Fig. 2. Illustration of the multiple connections to third parties (more than 50) generated in the background after visiting only 3 sites. The points where connections originate represent the websites while the little triangles represent the third parties contacted. This figure was obtained through the browser extension Disconnect [10].

By studying online tracking in this context, new findings are unveiled that are interesting with regard to user privacy risks.

II. BACKGROUND ON ONLINE TRACKING AND ADVERTISING

A. Online tracking

Online tracking refers to the activity of closely following a user wherever she "goes" while browsing the Web. This is possible because users leave innumerable footprints online, without even noticing it, when requesting for content to websites. IP address, operating system, browser type, plugins installed, patches applied, and browsing history are some examples of (context) information leaked in a single HTTP request. If aggregated and processed, said information could serve to build user profiles revealing location, shopping habits, entertainment preferences and even the gender of users.

The first potential tracker is thus the website (publisher) that the user visits. Thus, if tracking is performed from the publisher, it is called *first-party* tracking. In general, the audiences of first parties are pretty segmented, so the user tracking they might perform is usually innocuous. Some exceptions are the 'walled gardens' built by the Internet giants (e.g., Facebook), which concentrate services for millions of users within a single ecosystem.

Furthermore, a single user web request commonly triggers connections from the user browser to several *third parties* that receive part of the aforementioned contextual information. This information is used by third-parties to support real-time services such as personalized advertising or other services for websites, e.g., media hosting (by content distribution networks), load balancing, or social networking. Figure 1 illustrates the interactions triggered by a user browser request, which enable first and third-party tracking.

Undoubtedly, better online services are provided thanks to personalization and outsourcing; however, third party tracking supports the massive aggregation of user information (collected along multiple sites along the Web) in the hands of anyone aiming at paying for it [7]. This inevitably raises serious concerns regarding the privacy of users, because their online activity is received and processed by several entities

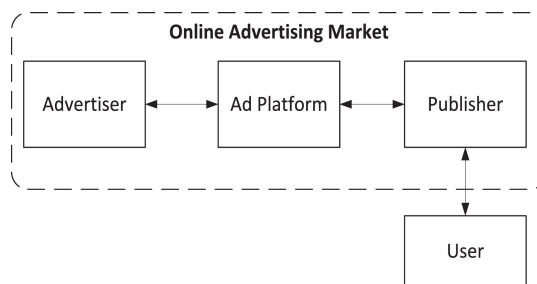


Fig. 3. Main components of the online advertising ecosystem.

that users had never heard of. Figure 2 illustrates the large number of connections to third parties (information flows) derived when a user visits only three websites.

B. Online advertising

Online advertising consists in displaying ads to users while they browse the Web. Many of the messages transmitted through advertising are (or are aimed to be) personalized. Four main players are involved in this service. First, *advertisers* are interested in paying for showing ads to ultimately sell a product or promote a brand. On the other hand, *publishers*, mainly websites, produce interesting content that attracts readers, i.e., potential customers; publishers are willing to sell ad space in their websites to display advertising material (ads from advertisers). *Ad platforms* are groups of entities that match the demand (from advertisers) and supply (from publishers) of online services. A marketplace is then built from ad platforms where ad spaces are automatically auctioned and delivered to the highest bidder in a fraction of a second. Given the complexity of buying and selling ad inventory under these conditions (automatically and in real time), publishers and advertisers trade in online platforms through other specialized entities: demand-side and supply-side platforms.

Finally, users are the agents whose single request to a website generates a chain of interactions among the entities aforementioned, ending, e.g., on a displayed ad. Figure 3 provides an overview of this ecosystem whose internal operation is described below.

In brief, this process may be depicted as follows. When a user browses a website engaged with an ad platform (*user impression*), not only an HTTP request is generated to the website. Through a piece of code sent to the browser in an HTTP reply, an automatic connection to the ad platform is triggered from the user's browser. Said third-party connection (ad call) requests the ad platform for ads to fill the ad spaces of the visited site.

Through a mechanism called *real-time bidding* (RTB), an ad platform auctions the user impression among the advertisers interested in displaying ads, awarding the ad space to the highest bidder. Advertisers bid decision is made based on user's metadata sent within the ad call. Personalization is here enforced by allowing advertisers to tailor their ads to the interests of users but also to their own strategies. In brief, ad platforms coordinate the roles of publishers, advertisers and users to maximize the resulting revenue.

C. Tracking tools for online advertising

Ad personalization requires tons of data about users. Also, to aggregate collected data on individual profiles, users have to be singled out during an impression. For this purpose, online tracking harnesses two main technologies: *cookie setting* and *fingerprinting*. Cookies are strings of text that a web server puts on the browser of a visiting user. In subsequent visits, the website retrieves the content of their cookies (usually a user ID) to recognize the visitor. A fingerprint, instead, is a string built from static characteristics of the applications and devices of a user (e.g., IP address, open ports, software versions, installed plugins). The combination of several of these data items could be pretty unique.

III. MEASUREMENT METHODOLOGY

In order to measure the extent of online tracking within the Ecuadorian context, automated visits are measured to the most popular Ecuadorian websites during December 2018. Upon each web request information about the requests spawn to third parties (i.e., to those different from the original destination) is captured.

This information was further processed to reveal the potential impact of third-party and advertising tracking on user privacy. This was first done by quantifying the presence of third-parties behind prominent Ecuadorian publishers. But we also tried to understand some of the dynamics behind the relationship between publishers and tracking entities, e.g., by unveiling the categories of the sites more prone to concentrate online tracking. Finally, other parameters related to cookie setting and transparency mechanisms were also processed to complement our vision of online tracking in Ecuador.

More details on the activities depicted above are given in the following subsections.

A. Crawling the most popular Ecuadorian websites

Online (third-party) tracking was triggered by generating first-party connections (HTTP requests) to several Ecuadorian sites. This implied a single visit to the home page of each

website with no further interactions. For this, we chose the most popular web pages to increase the probability of finding online ads that, evidently, motivate most of the tracking activities. Since generating visits and collecting the derived data imply a repetitive task for each website, we automatized these processes using the tool described below.

The list of URLs was extracted from the Alexa top 1 million site list (<https://www.alexa.com/>), from which 246 Ecuadorian websites were identified whose categories included news, sports, government, education, etc. For crawling and data collection, OpenWPM was used, a very versatile tool devoted to web measurement [1]. OpenWPM offers a programmable interface to orchestrate the main functions of a web browser, thus allowing automated web crawling and collection of tracking-related information (redirect, cookies and third-party calls) that is stored in a SQLite database.

Besides collecting third-party-related tracking, popular Ecuadorian sites were also crawled to examine the adoption of `ads.txt` [11], a project promoted by the Internet Advertising Bureau (IAB) to increase transparency in the programmatic advertising ecosystem and prevent fraud. It encourages publishers to publicly inform the companies they have authorized to sell their advertising inventory (ad spaces). Such publication is done through a text file (so much like the `robots.txt` standard) called `ads.txt` in the root context of the website. Interestingly, revealing such information could also serve as a transparency mechanism for users so we collected and processed the content of this file to confirm the results obtained when crawling third-party tracking.

B. Data Processing

To examine the magnitude of online tracking and the latent risks for privacy, the data obtained through web crawling was processed and the third-parties contacted (by publishers) and the cookies set by them were unveiled. The processing of this information consisted in filtering, aggregating and drawing the main results derived from an analysis that is divided as follows.

- Counting the (sometimes hundreds of) third-party requests triggered from each visit to popular Ecuadorian websites. Then, we identified the third parties (e.g., by domain) where our web requests were redirected. This was our first approach to measure the prevalence of online tracking since traffic to third parties commonly involves the disclosure of information that enables them to track end users. The degree of presence of these entities gave us a general idea of the resulting privacy risks.
- Some of the entities and traffic described above were related to not so intrusive services from the point of view of privacy (e.g., media hosting or load balancing). Thus, the same previous procedure was carried out but this time considering only the third-party tracking activities associated with online advertising (which is more privacy aggressive). To identify the advertising-related trackers, we applied the filtering rules (<https://easylist.to/>) of Ad-block Plus, a popular ad blocking browser extension, to the third-party URLs obtained above. By focusing on

TABLE I
PUBLISHERS THAT GENERATE THE GREATEST NUMBER OF THIRD-PARTY
REQUESTS FROM A SINGLE VISIT.

Publishers	# of third-party requests
studiofutbol.com.ec	1635
futbolecuador.com	1315
ecuadorinmediato.com	938
eldiario.ec	834
ecuagol.com	819
extra.ec	744
expreso.ec	632
derechoecuador.com	571
teamazonas.com	523
eluniverso.com	520

the online tracking prompted by advertising, we aimed at having a more realistic perspective of the impact on privacy.

- Examining the presence of third-party trackers aggregated by publisher and by publisher category. This approach gave us some clues to figure out the motivation of advertisers and online trackers to choose a specific website from which to track users in the Ecuadorian context.
- Interestingly, we found Ecuadorian government sites spawning web requests to known advertising and social media platforms. Thus, the analysis was focused on these websites to further research on the reason why these institutions were facilitating such third-party traffic.
- Finally, the cookies set by third-parties were analyzed to determine whether the corresponding strings could be considered as identifiers in the process of tracking users.

IV. RESULTS

A. Third-party tracking

Following the methodology described in Section III, we registered *hundreds* of third party requests as a result of a single interaction with Ecuadorian websites. Some of these requests are due to outsourced services that websites contract, e.g., for content distribution or load balancing. In any case, this implies a significant flow of information directly derived from user web requests. Thus, it is very likely that user information is being disclosed.

Beyond the privacy concerns raised by the leakage of personal information, the additional traffic towards entities different from the publisher may entail significant *monetary costs*, especially for users of mobile devices. This is a major issue when it was found that a single web request is followed by 57 additional requests on average, which is a huge amount of traffic not explicitly generated by the user.

In Table I, the publishers responsible for the greatest number of third-party requests have been outlined. According to the figure, *sports* and *news* websites are the ones generating the largest amount of traffic to third parties.

Many of these third-party requests could have the same destination, so it is convenient to identify the recipient entities (third-party trackers) by filtering the domain names from their destination URLs. As noted above, insofar these entities

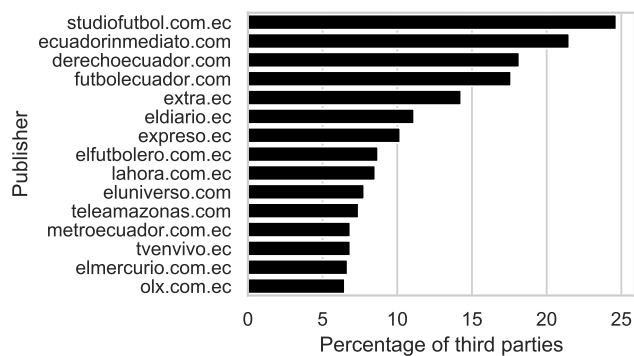


Fig. 4. Percentage of *third parties* associated with Ecuadorian publishers. Here the 15 websites with the greatest amount of third-parties are depicted.

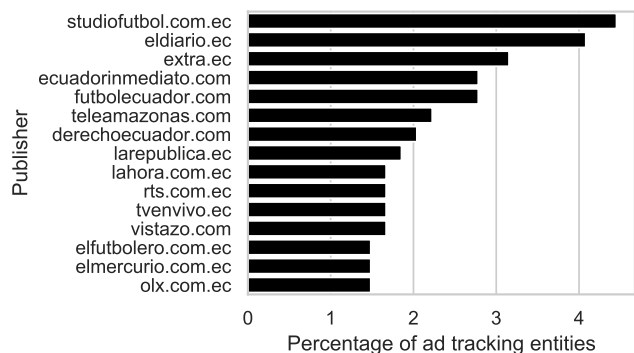


Fig. 5. Percentage of *ad-related* third parties associated with Ecuadorian publishers (with respect to the total amount of third parties found). 15 websites are depicted that contact the greatest quantity of third-parties.

receive so much indirect user traffic, they might become privacy attackers.

We found a grand total of 539 third-party entities present along Ecuadorian publishers, distributed in different proportions. Figure 4 depicts the 15 publishers that spawned most third-party traffic during our experiments. A sports site (*studio futbol.com.ec*) and a news site (*ecuadorinmediato.com*) redirect traffic to more than 20% of *all* the third party entities found in Ecuador. In fact, most of the publishers in the figure are again related to sports (particularly soccer), news and media. Although these findings could give us some intuition with regard to the concentration of online tracking, in the next section, where only the tracking due to advertising is considered, this is further analyzed.

B. Ad-related tracking

The tracking supporting online advertising is by far more intrusive in terms of privacy because it explicitly enables the collection of user data for personalization. As described in Section III, to measure such tracking, we filtered the destination URLs from our data set using the Easy list blocking rules. These rules are commonly used by popular web browser extensions to identify and block ad-related traffic. 147 entities were found receiving this traffic.

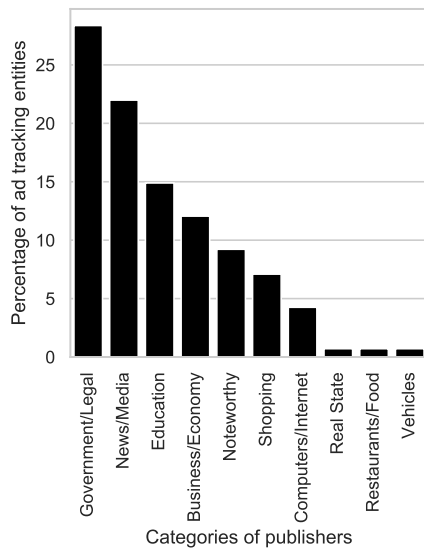


Fig. 6. Categories of Ecuadorian publishers where ad-related traffic is originated.

In Figure 5, the prevalence of ad-based trackers in publishers is plotted. The percentage of trackers in websites is evidently lower but the distribution looks very similar to that of Figure 4. Once again, sites categorized as News/Media are among the publishers more tracked by advertising entities. In general, this (including a few sports) sites are the preferred by advertisers that certainly look for wide audiences. Thus, the inherent tracking in such kind of publishers also grows.

To better understand how ad tracking is deployed along publishers, we categorized each of them and then aggregated the ad-related trackers according to such categories. This is shown in Figure 6. Around 20% of the ad tracking entities from our data set are present in News/Media sites. This means that several websites belonging to this category are engaged with ad-related tracking.

Interestingly enough, there was another category grouping several Ecuadorian sites, Government/Legal, where advertising-related tracking was prevalent. A lot of government websites were found that were triggering third-party requests to advertising platforms. Although the volume of third parties contacted was small, compared with News/Media, the amount of publishers involved was quite significant as depicted in Figure 6. This is further analyzed in the next subsection.

In general, categories shown in Figure 6 suggest, as expected, that ad-related online tracking is tightly related to publishers that offer mass consumption content (e.g., newspapers) and sales channels (e-commerce).

Finally, when aggregating the publishers of our data set according to the ad platform contacted, a remarkable concentration on Google-owned domains was found that exacerbates user privacy risks. As shown in Figure 7, at least seven domains associated with this company were receiving redirected tracking traffic from a lot of Ecuadorian publishers. In fact,

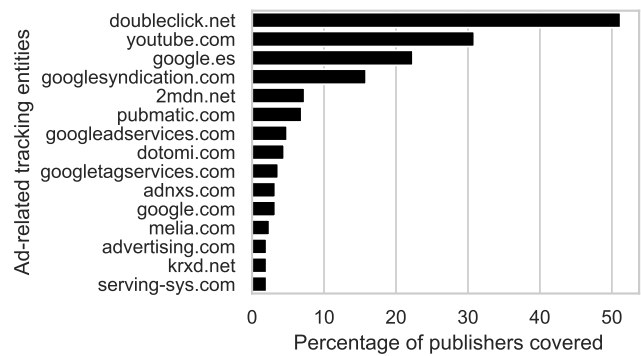


Fig. 7. Prevalence of *ad-related* tracking entities along the publishers examined. Google-owned domains (particularly `doubleclick.net`) appear on more than 50% of the publishers tested.

`doubleclick.net` collects requests from more than 50% of the publishers at hand.

C. Ad-related tracking on government websites

The presence of ad tracking in government sites was quite weird since they did not embed commercial advertising. In this regard, such third-party traffic was generated by a WordPress plugin installed by default. WordPress is a content management system, widely used by Ecuadorian websites, which simplifies the deployment of web pages but comes with a lot of modules activated by default, increasing the risk of security bugs. Through a new crawling, we found that 55.9% of government websites were developed on a WordPress template.

D. Specific privacy leaks

With the aim of identifying particular privacy leaks, we looked for *mouse tracking* entities. Currently, there are a lot of companies offering mouse analytics to capture the user experience with a website. However, privacy issues arise from the sensitive information that could be inferred from such biometrics [2]. The data collected was filtered by some keywords related to mouse tracking and found only 4 of such entities and 17 publishers using their services. `hotjar.com` was the most used (by 13 publishers).

Being cookies one of the pillars of online tracking, we were also interested in understanding their usage as user identifiers. Our approach was cataloging as an identifier each string that was more than six characters long. Thus, we found that 66% of the cookies set by third parties could have been considered as identifiers. Furthermore, unsurprisingly, hundreds of cookies had been set in at least 10 of the most popular Ecuadorian websites as depicted in Table II.

E. `ads.txt` as an ad transparency mechanism

Inspired by the work in [3], we examined the adoption of the `ads.txt` standard in the Ecuadorian context. For this, the selected publishers were crawled and, if available, the `ads.txt` file was collected and processed.

TABLE II
NUMBER OF COOKIES SET BY THIRD-PARTY TRAFFIC SPAWNED BY
POPULAR ECUADORIAN WEBSITES.

Publishers	# of cookies
studiofutbol.com.ec	348
teamazonas.com	343
futbolecuador.com	203
derechoecuador.com	199
eluniverso.com	151
ef.com.ec	139
olx.com.ec	119
metroecuador.com.ec	112
educarecuador.gob.ec	109
ecuadorinmediato.com	103
solidario.fin.ec	103

Around 15% of the popular publishers analyzed had adopted this transparency mechanism, even though Google had strongly encouraged its implementation at that time. This just revealed a still immature advertising market in this country.

When processing these `ads.txt` files, we found that four sites had more than 500 records authorizing to sell their ad spaces to several third-parties. This evidently implies a very high willingness to interact with ad platforms, with the privacy risks that it entails. Trough this analysis, we corroborated our finding that Google offered by far the most prevalent advertising platform.

V. CONCLUSIONS

To fuel digital advertising's automatic processes, online tracking enables the collection of several personal data items, e.g., IP addresses, location, timestamps, browsing history, language, etc. Through identification and external information, such items may significantly facilitate the work of third parties to unveil sensitive attributes about individuals. Online tracking in the Ecuadorian context is concentrated on a few popular News/Media publishers, mainly due to an underdeveloped online advertising market. This issue is evidenced by the lack of interest of publishers in adopting advertising industry standards such as `ads.txt` and the scarce presence of more specific tracking technologies such as mouse analytics. Anyhow, online tracking is still pervasive in this country: a large number of third parties are indirectly contacted by users, and a lot of tracking cookies set on the user side after a single web request. Such pervasiveness, mainly encouraged by advertising, the disclosure of granular user information and the need to meet real-time requirements, raise more privacy risks [12], [13]. Moreover, when ads are displayed, a few thousands of third parties receive metadata about millions of users for free, implying the establishment of a massive surveillance ecosystem [14]. Sadly, in a country as small as Ecuador, a thinner margin for anonymity is left. As this were not enough, Google, might be collecting user-related information from more than 60% of the publishers analyzed here. This privacy risk is exacerbated by operational issues of Ecuadorian websites, which might be provoking the leakage of user information (jeopardizing privacy) even when they are not engaged with advertising platforms.

Sadly, unlike many other countries, Ecuador does not have laws for privacy protection, so covering the specific concerns raised by online tracking and advertising, it seems, will remain a pending matter for long time. Future work in this context might involve, e.g., measuring the Ecuadorian user perception regarding this tracking behind the scenes, studying publishers' privacy policies towards advertising and third-party tracking, and even perform a comparative study of the situation at different countries.

ACKNOWLEDGMENT

This work was partly supported by the Spanish Ministry of Economy and Competitiveness (MINECO) through the project "MAGOS", ref. TEC2017-84197-C4-3-R. J. Parra-Arnau was supported by the Spanish government under grant TIN2016-80250-R and by the Catalan government under grant 2017 SGR 00705 and is currently the recipient of a Juan de la Cierva postdoctoral fellowship, IICI-2016-28239, from the Spanish Ministry of Economy and Competitiveness.

REFERENCES

- [1] S. Englehardt, A. Narayanan, "Online tracking: A 1-million-site measurement and analysis", In Proceedings of ACM CCS 2016, ACM, 2016.
- [2] P.E. Stillman, X. Shen, M. J. Ferguson, "How mouse-tracking can advance social cognitive theory", Trends in cognitive sciences, vol. 22, n. 6, pp. 531-543, 2018.
- [3] L. Olejnik, "Enhancing user transparency in online ads ecosystem with site self-disclosures", 2019, [Online]. Available: <https://lukaszolejnik.com/adstxt-transparency.pdf>. [Accessed: 15- May- 2019].
- [4] M. Graham, "Digital ad revenue in the US surpassed \$100 billion for the first time in 2018", CNBC, 2019. [Online]. Available: <https://www.cnbc.com/2019/05/07/digital-ad-revenue-in-the-us-topped-100-billion-for-the-first-time.html> [Accessed: 21- Jul- 2019].
- [5] C. Gayomali, "It Would Cost Each User \$232 A Year For An Ad-Free Internet, Study Finds", Fast Company, 2014. [Online]. Available: <https://www.fastcompany.com/3034670/it-would-cost-each-user-232-a-year-for-an-ad-free-internet-study-finds> [Accessed: 21- Jul- 2019].
- [6] A. C. Ecuador, "Constitución de la República del Ecuador", Tribunal Constitucional del Ecuador, Registro oficial Nro 449, 2008.
- [7] A. Rodríguez-Hoyos, J. Estrada-Jimenez, L. Urquiza-Aguilar, J. Parra-Arnau, J. Forné, "Digital hyper-transparency: leading e-government against privacy", Proc. Int. Conf. eDemocracy & eGovernment (ICEDEG), IEEE, p. 263-268, 2018.
- [8] SEPS, "Ley Orgánica de Transparencia y Acceso a la Información Pública", Superintendencia de Economía Popular y Solidaria, 2015. [Online]. Available: <http://www.seps.gob.ec/interna-npe?775>. [Accessed: 7- Oct- 2017].
- [9] S. Englehardt and A. Narayanan, "Online tracking: A 1-million-site measurement and analysis", Proc. ACM SIGSAC Conf. Comp. Commun. Secur., pp. 1388-1401, 2016.
- [10] Disconnect.me, "Disconnect — Take back your privacy", 2019. [Online]. Available: <https://disconnect.me/>. [Accessed: 23- Jul- 2019].
- [11] IAB, "ADS.TXT Authorized Digital Sellers", 2019. [Online]. Available: <https://iabtechlab.com/ads-txt/>. [Accessed: 23- Jul- 2019].
- [12] J. Achara, J. Parra-Arnau, C. Castelluccia, "MyTrackingChoices: Pacifying the Ad-Block War by Enforcing User Privacy Preferences", Proc. Annual Workshop Economics Inform. Secur., Berkeley, USA, June 2016.
- [13] S. Puglisi, D. Rebollo-Monedero, J. Forné, "On web user tracking of browsing patterns for personalised advertising", Intern. Journal Parallel, Emergent Distrib. Syst., vol. 32, no. 5, pp. 502-521, 2017.
- [14] J. Estrada-Jiménez, J. Parra-Arnau, A. Rodríguez-Hoyos, J. Forné, "On the regulation of personal data distribution in online advertising platforms", Engineering Applications of Artificial Intelligence, vol. 82, pp. 13-29, 2019.