Opportunities and Challenges of Ad-based Measurements from the Edge of the Network

Patricia Callejo[1,2] Conor Kelton[3] Narseo Vallina-Rodriguez[1,4] Rubén Cuevas[2]
Oliver Gasser[5] Christian Kreibich[4,6] Florian Wohlfart[5] Ángel Cuevas[2]

[1] IMDEA Networks Institute, [2] Universidad Carlos III Madrid, [3] Stony Brook University,
[4] ICSI, [5] Technical University of Munich, [6] Corelight

ABSTRACT

For many years, the research community, practitioners, and regulators have used myriad methods and tools to understand the complex structure and behavior of ISPs from the edge of the network. Unfortunately, the nature of these techniques forces the researcher to find a balance between ISP-coverage, user scale, and accuracy. In this paper we present AdTag, a network measurement paradigm that leverages the opportunistic nature of online targeted advertising to measure the Internet from the edge of the network. We discuss and formalize AdTag's design space—including technical, ethical, deployability and economic factors—and its potential to analyze a wide spectrum of Internet connectivity aspects from the browser. We run several experiments to demonstrate that AdTag can be tailored towards geographic and device-based user groups, finding also several challenges to be faced in order to maximize the number of samples. In a 7-day campaign, AdTag could access more than 20K ISPs at a global scale (185 countries) using millions of edge nodes.

CCS CONCEPTS

• **Networks** → **Network performance evaluation**; **Network performance analysis**; **Network measurement**;

KEYWORDS

Internet measurements, Advertising measurements

HotNets-XVI, November 30-December 1, 2017, Palo Alto, CA, USA

ACM ISBN 978-1-4503-5569-8/17/11...$15.00
https://doi.org/10.1145/3152434.3152895

1 INTRODUCTION

Tens of thousands of Internet Service Providers (ISP) offer Internet access to billions of customers from all over the world [18]. The Quality of Experience (QoE) perceived by Internet users is defined by myriad factors relating to the ISPs' network design, regulatory policies, network configuration, and operational decisions. In addition, a large number of research studies have revealed application-level and end-to-end connectivity violations, including traffic discrimination and network neutrality infringements [23, 24], DNS manipulations for profit [45], in-path TLS proxies [28], and traffic manipulation by in-path proxies [44], for example via HTTP header injection to facilitate advertising and user-tracking [41].

Revealing these manipulations, as well as identifying the culprits, is of significant interest to researchers, regulators, and end-users alike. This has motivated both the research community and practitioners to design and deploy tools to perform network measurements from the edge of the network. The resulting tools leverage dedicated testbeds, crowdsourced measurements, and VPN services to gather insights into the edge view of the network. While powerful, they all possess inherent drawbacks such as limited geographical and ISP coverage, or short-term experiment lifespan.

Despite years of network measurement and other studies conducted from the edge of the network, pervasive access to the network edge in order to facilitate measurements has remained elusive. To close this gap, we propose AdTag, an approach that leverages online advertising to launch network measurements at a global scale, in a time- and cost-effective manner. The nature of online advertising services make them an ideal, yet underused, distribution channel for launching rich network measurements either globally, opportunistically or focused on specific regions using the targeting mechanisms provided by online advertising service.

While select previous studies have likewise employed ad-driven measurements, they have done so in an opportunistic, experiment-specific manner. Our goal in this work is to take a step back and consider the experimental apparatus of JavaScript-enabled ad placement

Project	Nodes†/IPs*	ASes	Countries	Time	Deployment strategy
AdTag	2,500,000*	20,700	185	7 days	Targeted ads
RIPE Atlas	9,300†	3,300	181	6 years	Testbed / Dedicated node
Archipelago	181†	146	60	10 years	Testbed / Dedicated node
Netalyzr	2,200,000*	14,500	196	6 years	Crowdsourcing / Mobile app, browser applet
Luminati	1,300,000*	14,700	172	5 days	P2P-based VPNs

Table 1: Comparison of a global AdTag campaign with previous studies in terms of network coverage, measurement duration, and deployment strategy. (*: number of sessions; †: number of nodes)

and explore its broader feasibility for network measurement. We discuss aspects and challenges inherent to the distribution channel (*i.e.,* ad networks), the execution environment (*i.e.,* the browser), ethical concerns, and possible legal constraints. We demonstrate that AdTag provides a viable and promising alternative platform for conducting a wide range of network measurements at scale, driven by web-based JavaScript APIs.

2 BACKGROUND

Existing edge-driven measurement techniques fall into four broad categories that we survey in this section. Table 1 summarizes our findings.

Dedicated testbeds: Several dedicated measurement testbeds exist. RIPE Atlas [35], CAIDA's Archipelago (Ark) Measurements Infrastructure [5], the MONROE Mobile Broadband measurements platform [25], BISmark [2], and PlanetLab [31] are prominent examples. RIPE Atlas, Ark, and BISmark require dedicated hardware typically hosted by volunteers or academic institutions. As a result, these platforms typically possess limited geographical and ISP coverage due to their high deployment cost. Moreover, these platforms differ widely in openness and the types of tests one can execute.

Crowdsourcing: Researchers have developed several user-friendly tools to help users to understand the behavior of their network. In exchange, the research teams collect valuable, oftentimes anonymized, real-world data about the access link. Examples include the ICSI Netalyzr [19], DASU [37], MobiPerf [36], and Encore [4], which embeds JavaScript code on popular landing pages, unbeknownst to users. These tools are available as apps for mobile devices, browser-based clients, command line clients, or plugins for BitTorrent clients. As opposed to measurements run on dedicated testbeds, measurement campaigns following a crowd-sourcing strategy allow researchers to maximize ISP and user coverage without necessarily sacrificing data accuracy and detail. Commercial products like Ookla's SpeedTest [29], and measurement campaigns run by regulators (*e.g.,* FCC's speedtest[10]) have also followed this model with great success. Unfortunately, the majority of these tools only provide a snapshot of the network at a given time when the user executes the tool. This limits their ability to run longitudinally, and to measure behavior at a point in time chosen by the researcher.

VPN-based studies: A number of research efforts have leveraged VPN services to penetrate ISPs all over the world. One popular VPN service used by researchers is Luminati [20], a commercial VPN service that provides vantage points in more than 20M residential and enterprise IPs. Luminati has been used to detect traffic manipulations inflicted by in-path HTTP proxies [39] and end-to-end violations in the Internet [6]. Further, Luminati's low-end monthly price is $500 for 40GB of traffic. However, recent studies have questioned the ethical, privacy and security aspects of such VPN services [16], and it is unclear whether the egress points can also actively manipulate user's traffic. Other projects like ICLab have also used commercial VPN services to conduct censorship analysis [33] at a global scale. Unfortunately, recent studies have questioned the ISP coverage of these services [43], which may bias the experimental results.

Targeted ads: Ads have rarely been used for academic Internet measurements on a large scale. O'Neill *et al.* leveraged Flash-based ads to identify the presence of TLS proxies [28]. Since most modern browsers and ad networks move to deprecate or disable Flash, [12] it no longer offers a sustainable deployment mechanism. The same holds true for Java applets. Geoff Huston used advertising campaigns for APNIC Labs' IPv6 Measurement System [15], achieving good coverage by downloading a tracking pixel using JavaScript and Flash ads. A recent paper by Corner *et al.* proposes advertisement as a platform for large-scale network measurements. The authors demonstrate its ability to improve geo-IP databases, conduct bandwidth measurement and the identifiability of mobile users [7]. It corroborates our proposal of an advertisement-driven solution to edge measurement, but their study is focused solely on mobile measurements, namely device battery management and GeoIP databases.

3 ADTAG

AdTag leverages ad networks for conducting network measurements at a global scale, in a time- and cost-effective manner. However, distributing complex network measurements through ad networks and running them

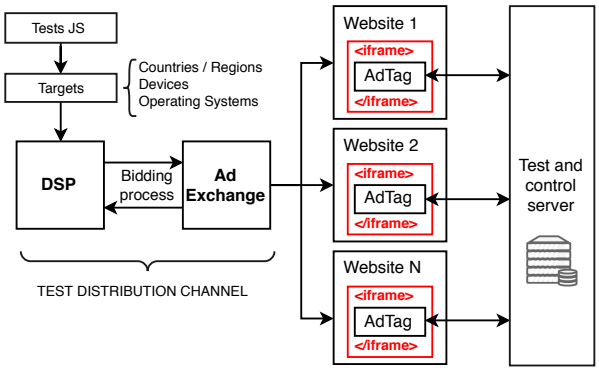

Figure 1: AdTag architecture, distribution channel and client-server components for measurements.

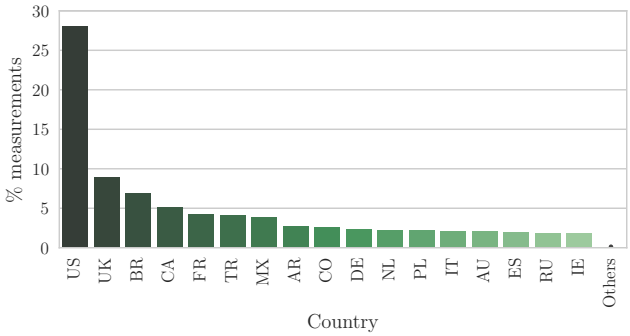

Figure 2: Distribution of user IPs around the world.

on the browser poses several challenges which have not been systematically studied so far.

In this section we discuss AdTag's design space[1]. First, we describe our test distribution channels through ad networks. Then, we put our focus on understanding aspects inherent to ad networks such as the cost of launching campaigns, our ability to target specific user groups and platforms, the available execution window, and ethical aspects. For these, we use empirical data that we obtained from a purposely-run advertising campaign launched through a Demand Side Platform (DSP). We also give background on the operation of such platforms.

3.1 Deploying Network Measurements

We deploy AdTag measurements using real advertising campaigns configured through a Demand Side Platform (DSP). Briefly explained, the current online advertising ecosystem [3], typically called *programmatic advertising*, is a complex one, composed by multiple intermediaries. The ad spaces available in a publisher website are typically handled by ad networks or Supply Side Platforms (SSPs), those intermediaries are in charge of selling the ad spaces. From the buying side, the advertisers typically rely on agencies or DSPs to manage their campaigns. A DSP is an intermediary platform providing advertisers unified access to multiple vendors (Ad Networks and Ad Exchanges), each selling ad spaces from a pool of websites and mobile apps. It also enables advertisers to configure targeting parameters for their campaigns (geographical location, device type, etc).

As a proof of concept, we run a 7-day campaign using 9 of the more than 20 ad networks provided by a DSP.[2] This campaign provides us with more than 3M measurements from 2.5M unique IP addresses covering 185 different countries. This rivals the number of sessions initiated by the crowd-sourced Netalyzr platform over a timespan of 6 years, underscoring the method's broad reach. AdTag leverages HTML5-based ads [11, 17] to execute JavaScript-based active network measurements from the edge of the network. JavaScript (JS) allows us to embed different pieces of code to conduct a wide range of network measurements, which will be distributed at a global scale through advertising campaigns as illustrated in Figure 1. AdTag is constrained to the features and APIs provided by end-user browsers. It is important to remark that the DSP renders the ads in an iFrame, which sandboxes the JS code. This prevents it from interacting directly with the parent window, including via cookies. Apart from those constraints, our DSP enables us to perform all the measurements explained in this paper. Note that others limitations may apply depending on the DSP.

3.2 Targeting ISPs and Locations

Targeting measurements to specific ISPs and geographical locations allow researchers to precisely analyze and penetrate particular providers. This ability is determined by the accuracy of the targeting mechanisms provided by the DSP. Most DSPs allow targeting campaigns based on location, device type (*e.g.,* desktop vs. mobile), and even operating system. We use this feature to configure the campaigns to the experiment's needs and to target specific ISPs.

We perform several experiments to analyze the feasibility of targeting ISPs and platforms, and evaluate the precision of the DSP's target mechanisms. We use MaxMind's database [22] to geolocate client IP addresses. While research has shown that the use of IP geolocation databases can introduce biases [32], we believe them still to be indicative of the overall deployment. Our global ad campaign covers 185 countries, with the majority of measurements coming from clients in the US (28%), UK (8.8%), Brazil (6.8%), and Canada (5.1%). Figure 2 shows the overall geographical coverage obtained with our global campaign, which covers 185 different countries.

When geolocating US-based IP addresses, we can see in Table 2 that most of the impressions come from large

[1]The online advertising industry uses the term *ad tag* to refer to a piece of code typically used to monitor ad behavior.
[2]By request of the DSP we cannot share its name.

ISP Name	# samples	% samples
Comcast Cable	149.4K	17.7
CenturyLink	99.3K	11.7
Time Warner Cable	85.8K	10.1
AT&T U-verse	69.3K	8.2
Cox Communications	37.8K	4.5

Table 2: Top 5 most representative ISPs from the USA according to the results of our global campaign.

fixed-line and mobile ISPs like Comcast. However, our advertising campaign also allows accessing a fair number of small ISPs such as NTS Connections (AS46698) and Northwest Open Access Network (AS16713), both of them with at least one hundred samples.

These observed coverage distributions are expected as we did not use precise geographical targeting and thus received biases in impressions towards the US, where most of the websites are hosted, and towards ISPs with a large customer base. To target a particular ISP, the researcher can adapt the campaign using various features offered by DSPs. Some DSPs allow deployment on a country or city-level, this can be used to target the area where a desired ISP is known to operate, maximizing the number of valid samples. Other DSPs allow researchers act similarly by specifying IP ranges for deployment [7].

To validate these proposed solutions, we ran two 1-day 50K sample experiments targeting the USA and NYC, respectively. In the country-level experiment, 97% of the users had a US-based IP address. The rest of the samples came from a handful of countries, namely Canada (2% of the total samples). The results of the city-level experiment show similar accuracy.

3.3 Price

Running online advertising campaigns comes at a cost. However, it is possible to leverage different strategies to maximize the geographical coverage while keeping the budget under control. For instance, in our global campaign we fixed the CPM (Cost Per Mille) budget. Our DSP allows CPMs starting at $0.10. Therefore, it is possible to launch campaigns at this minimum CPM cost and consider higher CPMs in order to increase geographical and ISP coverage when needed (for instance, to target under-represented geographical areas).

For the majority of network measurements, user clicks are irrelevant. User interaction may be only needed when their feedback is required, as in the case of QoE experiments. As a result, AdTag does not need to apply any campaign optimization based on CPC (Cost per Click), notably reducing the budget requirements to launch measurement campaigns.

As AdTag is running on a large number of heterogeneous systems and configurations, our measurements are subject to multiple sources of errors which can cause data loss, such as browser extensions preventing JavaScript (*e.g.,* ad-blockers [21, 27]), transient network disruptions, and limited browser API support. Overall, comparing the DSP reports and the data gathered by AdTag, we witness 15% data loss on average per campaign.

	Percentiles		
Device	25^{th}	50^{th}	75^{th}
Mobile	7.8s	30.1s	105.9s (>1min)
Desktop	14.3s	33.6s	110.7s (>1min)

Table 3: Execution time percentiles per device type.

An estimation of the cost per campaign, assuming an average CPM of $0.10[3] and a conservative efficiency ratio of 80%, resulted in approximately 1M measurements for a $125 budget, more cost efficient than previous research driven by ad placements ($5K for almost 3M successful measurements) [28]. We conclude that running measurements using online ads is 1) more flexible, 2) increases ISP coverage, and 3) is more economic than using VPN-based systems.

3.4 Execution Window

A website—including any embedded element, such as ads—may be active in the browser for only a short period of time: if the user opens a new website or simply closes the tab, the JavaScript code running AdTag tests will be immediately interrupted. As a result, it is important to know for how long we can run our measurements, *i.e.,* the *execution window*.

We use the data provided by our global campaign to estimate the expected execution window. Our results suggest that 75% of ads are active for more than $11s$, regardless of end-user platform, with a median time of $33s$. Table 3 shows the 25^{th}, 50^{th} and 75^{th} percentiles of the execution window for desktop and mobile devices. We see significant differences in the execution window depending on the platform: 75% of ads rendered on the desktop are active for at least $15s$ whereas this decreases to just $8s$ for mobile devices. This analysis suggests that being time-conscious is critical to the experiment's design. Tests should launch and complete quickly, and should be scheduled opportunistically to make use of long-running ad displays.

3.5 Ethical and Legal Considerations

In contrast to many crowdsourced measurement techniques, ad-driven measurements such as AdTag's are likely to run *unbeknownst* to the crowd participant, placing particular responsibility on the measurement orchestrator. While previous work has demonstrated the community's sensitivity to this type of experiment [4], recent work continues to operate in similar fashion [7, 42]. In the following we review ethical and legal aspects of this responsibility, and position our work in this context.

[3]We pay the minimum CPM allowed by our DSP, as our goal is maximizing the number of impressions and not their quality.

We acknowledge and remind experimenters that ad-driven measurements bear the potential of harm to the client. Consider an experiment that collects client IP addresses together with HTTP request headers and their potential to profile individual users. While the ethical sensitivity of such experiments is clear, the experimenter also needs to be cognizant of potential legal constraints of the measurement, such as when an ad connects to websites deemed illegal in the user's country, or the local jurisdiction considers the collected information personally identifiable. Our work on AdTag has not and will not engage in practices that violate these concerns, and we obtained ethics approval from IMDEA's ethics review board before conducting the experiments. [4] We also followed the ethical guidelines defined by the community [8, 30].

In the context of ad-based measurements informed consent [8] is difficult to obtain. The option of using ad-blocking software only offers blunt control over ad displays, and while the `Do Not Track` request header could serve as a possible signal to the experimenter, its applicability to arbitrary measurements remains unclear to both users and experimenters. Accordingly, we did not obtain informed consent from AdTag's participants. For this specific paper, we did not collect any personal or sensitive information from the user, anonymizing collected data. The ads rendered in our campaign pointed to one of our ongoing research projects [40], ensuring that all connections were made to a safe and uncensored server under our control. Finally, to the best of our knowledge, our tests also comply with the terms of use of our chosen DSP.

4 NETWORK MEASUREMENTS IN THE BROWSER

Modern web browsers run powerful JS engines that offer a rich suite of networking libraries to web developers. Many of the client-side APIs used in AdTag have been standardized by the web community:

XMLHttpRequest (XHR): This API allows clients to communicate to servers via HTTP(s) protocols, allowing custom crafted methods, headers, and payloads [13].

WebSocket: This standard allows delivering custom application-level data in a bi-directional manner between a client browser and a server over TCP [14].

Network Information API: Most DSPs claim to be able to run ad campaigns restricted to mobile devices. However, mobile devices may not necessarily be connected over a cellular link: users can also access the Internet from their smartphones over WiFi. AdTag can use the Network Information API [9] supported by Firefox and Chrome browsers on Android to obtain ground-truth about the access link technology of the device.

WebRTC: This API, not completely standardized yet by the W3C but already fully supported by most browsers [1], allows communicating custom application data (namely for video and audio) over a bi-directional UDP channel. WebRTC also provides access to many of the utilities required for establishing peer to peer connections, including methods to perform NAT traversal.

As opposed to programming languages with a full network stack like Java and Flash, JS networking APIs have several technical constraints that limit our ability to implement certain network measurements. Restrictions on WebSocket and WebRTC do not allow the creation of data directly over TCP/UDP such that they could be used to exactly mimic and modify existing application-level protocols. Even though a WebSocket can carry arbitrary unencrypted data over TCP, it requires a connection phase between client and server using HTTP(s) before proceeding with any data transfer. It also has its own custom headers, which encapsulate the data. WebRTC UDP is restricted in a similar manner, requiring DTLS encryption for any data channel and encapsulating the data channel within SCTP. As a result, AdTag will not be able to directly test certain UDP-based protocols and Internet sub-systems like DNS [45].

Nevertheless, the implications of what these APIs allow in terms of network measurements are still enormous as we will demonstrate in Section 4.2. As UDP traffic via WebRTC is delivered over SCTP at the application level, it provides a good balance between accuracy and efficiency for network measurements. This allows to choose whether SCTP data is guaranteed to be delivered in order, reliably, neither, or both. Consequently, performance reliant tests, such as latency or timeout tests, can be more accurate than those done over TCP, where overheads occur due to mandatory inclusion of reliable/in-order delivery and state maintaining.

Alternatively, tests where accuracy is the priority such as outbound port scans, can take advantage of the added utility of probing the lower levels with UDP flows while still producing reliable results at the application level.

[4] Specifically, AdTag's experiments were conducted exclusively by IMDEA Networks Institute personnel, whose ethics review board approved informally our study before we engaged our measurements, following the internal guidelines and procedures in place at the time of writing. IMDEA Networks' ethics review board is composed of three senior researchers. Due to the institute's small size no other more independent body is responsible for research ethics review. We will further strengthen our ethical review process by requesting written approval for future AdTag measurements. We recommend that other researchers also follow this approach for AdTag-type measurements due to the possible ethical implications described in this section.

4.1 Browser support

We instrumented our advertising campaign to measure browser's API support in the wild. Table 4 shows a breakdown of dominant browsers, according to their `User-Agent` field, that we identified during our global campaign, ordered by the percentage of successful measurements run on each one of them over the total. For each browser and JS API, we report the minimum version supporting a given API. `n/a` indicates that a given browser does not support such technology yet. The percentage value for each technology reports the percentage of users for a given browser running at least the minimum browser version supporting this technology.

45% of the most common browsers (shown in Table 4) of our global campaign, were launched on browsers supporting the three networking APIs simultaneously. As we can see, most measurements come from Chrome users, which guarantees that a large number of tests will be executed on browsers with full API support. The analysis also reveals that mobile browsers provide more limited APIs than their desktop counterparts. Unfortunately, DSPs do not allow targeting end users according to API support. Therefore, understanding browser API support is key to plan complex measurement campaigns and adjust their budget accordingly by predicting how many impressions will be required to obtain statistically representative results.

4.2 Use cases

JS libraries can be used to bootstrap a wide range of network measurements through AdTag. Some may require only instrumenting the client-side JS. However, others may require interaction between the client and collaborative server, as illustrated in Figure 1. Next, we present a non-exhaustive list of interesting network measurements—some based on previous measurement tools using full-stack programming languages—that can be successfully ported to JS.

Detecting middleboxes and traffic manipulation: A careful instrumentation of both the client- and the server-side of AdTag can reveal the presence of HTTP and HTTPS middleboxes and if they perform any traffic manipulation. Using the WebSocket and XHR libraries, we can force the client and the server to speak custom variants of HTTP over TCP, a technique proved valid to identify and characterize HTTP(s) proxies [19, 41, 44].

NAT detection and characterization: WebRTC allows performing STUN and TURN requests that can be used to study NATs at scale. In this case, a STUN/TURN server is required. Because of the direct access of the user to proper protocols over UDP for NAT traversal through STUN and TURN, the client can obtain data regarding its IP, probe for NAT existence, check for middlebox state and identify port allocation policies. These features were previously limited to Java-based frameworks like NAT-Analyzer [26] and Netalyzr [19, 34].

Browser	%	WebRTC		WebSocket		WebWorker	
		Ver.	%	Ver.	%	Ver.	%
Chrome	34.5	49	97	49.0	97.0	49.0	97.0
Mobile Safari	21.7	n/a		9.3	14.3	9.3	14.3
Chrome Mobile	19.8	59	56	59.0	56.0	59.0	56.0
Firefox	5.7	52	88	52.0	88.0	52.0	88.0
Safari	4.6	n/a		9.3	95.0	9.3	95.0

Table 4: Top 5 most common browsers in our global campaign and the minimum version supporting relevant JS APIs. The percentage value for each API is computed over the total number of browsers of a given kind.

CDN performance: CDN performance highly depends on the replica selection algorithm and DNS resolution. AdTag clients can fetch one (or more) small object(s) from a CDN provider hence providing detailed performance metrics such as the time-to-first byte (TTFB), and the location of the assigned replica.

IP classification: AdTag-based tests can help to classify a given IP address along different dimensions: by network type (*i.e.,* residential, enterprise or mobile) and characteristics (*e.g.,* proxied or NATed). The mapping of an IP to UAs reveals the sharing condition of an IP address. This can complement existing IP intelligence datasets, helping to further contextualize the data provided by IP blacklists, WHOIS records, and geo-IP services [38].

5 CONCLUSION

In this paper we have presented and discussed AdTag, a measurement platform that leverages online advertising to quickly conduct experiments at global scale. AdTag leverages ad networks' ability to target specific client populations in order to analyze the Internet from the edge of the network. We discussed AdTag's design space, including its ability to target specific networks and devices, typical campaign costs, ethical and legal concerns, as well as technical challenges imposed by browser runtimes. Common JavaScript APIs can serve to detect and characterize middleboxes such as proxies and NATs, analyze CDN performance, or furnish the input for IP address classification. Our empirical experiments placed ads in 9 ad networks and confirm the ability to target specific ISPs and geographic locations at low cost, facilitating large-scale data collection within days.

Acknowledgments: We thank our anonymous reviewers for their valuable feedback. This work is partially supported by the European Union through the H2020 TYPES (653449) and ReCRED (653417) projects, the Spanish Ministry of Economy and Competitiveness through the DRONEXT project (TEC 2014-54335-C4-2-R), the European Social Fund through Ramón y Cajal project RYC-2015-17732, the Madrid Community through the BRADE project (P2013/ICE-2958), and the German BMBF projects X-CHECK (16KIS0530) and DecADe (16KIS0538).

REFERENCES

[1] B. Aboba, C. Jennings, A. Narayanan, T. Brandstetter, D. Burnett, and A. Bergkvist. WebRTC 1.0: Real-time communication between browsers. W3C working draft, 2017.

[2] BISmark. Broadband Internet Service Benchmark. http://projectbismark.net, 2017.

[3] K. Bridge. Online advertising explained. http://www.kbridge.org/en/online-advertising-explained-dmps-ssps-dsps-and-rtb/, 2013.

[4] S. Burnett and N. Feamster. Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests. 2015.

[5] CAIDA. Archipelago (ARK) Measurements Infrastructure. http://www.caida.org/projects/ark/, 2017.

[6] T. Chung, D. Choffnes, and A. Mislove. Tunneling for Transparency: A Large-Scale Analysis of End-to-End Violations in the Internet. In *ACM IMC*, 2016.

[7] M. Corner, B. Levine, O. Ismail, and A. Upreti. Advertising-based Measurement: A Platform of 7 Billion Mobile Devices. *ACM Mobicom*, 2017.

[8] D. Dittrich and E. Kenneally. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research. *US DHS*, 2012.

[9] W. C. G. Draft. Network Information API. http://wicg.github.io/netinfo/, 2017.

[10] FCC. Measuring Broadband America. https://www.fcc.gov/general/measuring-broadband-america, 2017.

[11] Google. Build an HTML5 creative. https://support.google.com/richmedia/answer/2672542?hl=en, 2013.

[12] Google. Update your Flash ads. https://support.google.com/adwords/answer/6249073, 2017.

[13] S. Hallvord, A. Van Kesteren, J. Aubourg, and J. Song. XMLHttpRequest level 1. W3C note, 2016.

[14] I. Hickson. The WebSocket API. Candidate recommendation, 2012.

[15] G. Huston. APNIC Labs IPv6 Measurement System. https://labs.apnic.net/?p=348, 2013.

[16] M. Ikram, N. Vallina-Rodriguez, S. Seneviratne, M. A. Kaafar, and V. Paxson. An Analysis of the Privacy and Security Risks of Android VPN Permission-enabled Apps. In *ACM IMC*, 2016.

[17] Internet Advertising Bureau (IAB). HTML5 for Digital Advertising v2.0. https://www.iab.com/guidelines/html5-for-digital-advertising-guidance-for-ad-designers-creative-technologists, 2017.

[18] Internet Society. Global Internet Maps. http://www.internetsociety.org/map/global-internet-report, 2017.

[19] C. Kreibich, N. Weaver, B. Nechaev, and V. Paxson. Netalyzr: illuminating the edge network. In *ACM IMC*, 2010.

[20] Luminati. https://luminati.io, 2017.

[21] M. Malloy, M. McNamara, A. Cahn, and P. Barford. Ad blockers: Global prevalence and impact. In *ACM IMC*, 2016.

[22] MaxMind. GeoIP Database. https://www.maxmind.com, 2017.

[23] A. Molavi Kakhki, F. Li, D. Choffnes, E. Katz-Bassett, and A. Mislove. BingeOn Under the Microscope: Understanding T-Mobiles Zero-Rating Implementation. In *ACM Workshop on QoE-based Analysis and Management of Data Communication Networks*, 2016.

[24] A. Molavi Kakhki, A. Razaghpanah, A. Li, H. Koo, R. Golani, D. Choffnes, P. Gill, and A. Mislove. Identifying traffic differentiation in mobile networks. In *ACM IMC*, 2015.

[25] MONROE. Measuring Mobile Broadband Networks in Europe. https://www.monroe-project.eu, 2017.

[26] A. Müller, F. Wohlfart, and G. Carle. Analysis and Topology-based Traversal of Cascaded Large Scale NATs. In *ACM HotMiddlebox*, 2013.

[27] R. Nithyanand, S. Khattak, M. Javed, N. Vallina-Rodriguez, M. Falahrastegar, J. E. Powles, E. De Cristofaro, H. Haddadi, and S. Murdoch. Adblocking and counter blocking: A slice of the arms race. In *USENIX FOCI*, 2016.

[28] M. O'Neill, S. Ruoti, K. Seamons, and D. Zappala. TLS proxies: Friend or Foe? In *ACM IMC*, 2016.

[29] Ookla Speedtest. http://www.ookla.com, 2017.

[30] C. Partridge and M. Allman. Ethical considerations in network measurement papers. *Communications of the ACM*, 2016.

[31] PlanetLab. https://www.planet-lab.org, 2017.

[32] I. Poese, S. Uhlig, M. A. Kaafar, B. Donnet, and B. Gueye. Ip geolocation databases: Unreliable? 2011.

[33] A. Razaghpanah, A. Li, A. Filastò, R. Nithyanand, V. Ververis, W. Scott, and P. Gill. Exploring the design space of longitudinal censorship measurement platforms. *arXiv preprint*, 2016.

[34] P. Richter, F. Wohlfart, N. Vallina-Rodriguez, M. Allman, R. Bush, A. Feldmann, C. Kreibich, N. Weaver, and V. Paxson. A multi-perspective analysis of carrier-grade NAT deployment. In *ACM IMC*, 2016.

[35] RIPE. RIPE Atlas. https://atlas.ripe.net, 2017.

[36] S. Rosen, H. Yao, A. Nikravesh, Y. Jia, D. Choffnes, and M. Mao. Mapping global mobile performance trends with Mobilyzer and MobiPerf. In *ACM MobiSys*, 2014.

[37] M. A. Sánchez, J. S. Otto, Z. Bischof, D. Choffnes, F. Bustamante, B. Krishnamurthy, and W. Willinger. Dasu: Pushing Experiments to the Internet's Edge. In *USENIX NSDI*, 2013.

[38] Q. Scheitle, O. Gasser, P. Sattler, and G. Carle. HLOC: Hints-Based Geolocation Leveraging Multiple Measurement Frameworks. In *IEEE TMA*, 2017.

[39] G. Tyson, S. Huang, F. Cuadrado, I. Castro, V. C. Perta, A. Sathiaseelan, and S. Uhlig. Exploring http header manipulation in-the-wild. In *WWW*, 2017.

[40] UC3M. Facebook Data Validation Tool. https://fdvt.org, 2017.

[41] N. Vallina-Rodriguez, S. Sundaresan, C. Kreibich, and V. Paxson. Header Enrichment or ISP Enrichment? Emerging Privacy Threats in Mobile Networks. In *ACM HotMiddlebox*, 2015.

[42] P. Vines, F. Roesner, and T. Kohno. Exploring ADINT: Using Ad Targeting for Surveillance on a Budget — or — How Alice Can Buy Ads to Track Bob. In *ACM Workshop on Privacy in the Electronic Society*, 2017.

[43] VPNscam.com. How to avoid VPN scams in 2017-2018. http://vpnscam.com/how-to-avoid-vpn-scams-in-2017-2018/, 2017.

[44] N. Weaver, C. Kreibich, M. Dam, and V. Paxson. Here be web proxies. In *PAM*, 2014.

[45] N. Weaver, C. Kreibich, and V. Paxson. Redirecting DNS for Ads and Profit. In *USENIX FOCI*, 2011.