# Who's Most Targeted and Does My New Adblocker Really Help: A Profile-based Evaluation of Personalized Advertising

Sofia Bertmar Johanna Gerhardsen Alice Ekblad Anna Höglund Julia Mineur
Isabell Öknegård Enavall Minh-Ha Le Niklas Carlsson
Linköping University, Sweden

## ABSTRACT

There is limited prior work studying how the ad personalization experienced by different users is impacted by the use of adblockers, geographic location, the user's persona, or what browser they use. To address this void, this paper presents a novel profile-based evaluation of the personalization experienced by carefully crafted user profiles. Our evaluation framework impersonates different users and captures how the personalization changes over time, how it changes when adding or removing an extension, and perhaps most importantly how the results differ depending on the profile's persona (e.g., interest, occupation, age, gender), geographic location (US East, US West, UK), what browser extension they use (none, AdBlock, AdBlock Plus, Ghostery, CatBlock), what browser they use (Chrome, Firefox), and whether they are logged in to their Google account. By comparing and contrasting observed differences we provide insights that help explain why some user groups may feel more targeted than others and why some people may feel even more targeted after having turned on their adblocker.

## CCS CONCEPTS

**• Security and privacy → Privacy protections**; **• Information systems → Online advertising**.

## KEYWORDS

Profile-based analysis, Personalized ads, Targeted advertisement, Privacy, Browser extension

**ACM Reference Format:**

Sofia Bertmar Johanna Gerhardsen Alice Ekblad Anna Höglund Julia Mineur, Isabell Öknegård Enavall Minh-Ha Le Niklas Carlsson. 2021. Who's Most Targeted and Does My New Adblocker Really Help: A Profile-based Evaluation of Personalized Advertising. In *Proceedings of the 20th Workshop on Privacy in the Electronic Society (WPES '21), November 15, 2021, Virtual Event, Republic of Korea.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3463676.3485617

## 1 INTRODUCTION

The success stories of targeted and personalized advertisements can be intimidating and offend some. While some people have argued that the exposure of such ads is a price that users must pay to receive free content/service (e.g., [27]), others have buckled down and developed adblockers and other privacy enhancing browser extensions [3, 15, 20, 21]. Such extensions typically attempt to block third-party trackers, advertisements, or even replace the advertisements with an alternative image. Although the use of these services has their own privacy and security risks [10, 13], adblockers and other privacy enhancing browser extensions have become a popular way to reduce the number of ads that a user is exposed to.

This has prompted several papers to study the effectiveness of adblockers [18] and their performance tradeoffs [10]. However, to the best of our knowledge, no prior work has studied the personalization experienced by realistic users selecting to use these technologies and how the level of personalization is affected by various other factors. In this work, we present a novel profile-based evaluation of the personalization experienced by carefully crafted user profiles that aim to surf the web in a similar fashion as many modern web users. We next summarize our key contributions.

First, we develop a Selenium-based data-collection tool that emulates user sessions of specific personas. Second, using the tool, we implement an experimental design that allows us to evaluate the impact that many different factors have on the personalization perceived by six (or in some cases three) carefully handcrafted user profiles, while controlling for the other factors of consideration. Here, we study how the personalization changes over time for users with different persona (e.g., interest, occupation, age, gender), how the results differ depending on the profile's geographic location (US East, US West, UK), what browser extension they use (none, AdBlock [5], AdBlock Plus [26], Ghostery [2], CatBlock [1]), what browser they use (Chrome, Firefox), and whether they are logged in to their Google account or not.

Third, we performed a longitudinal measurement campaign for 21 days. Given the above experimental design, we ran 51 VMs in parallel for 21 days, where we used the first 14 days as our main experiments (described above) and used the last seven days to study the effects of adding or removing an extension. In total, the users visited 178,650 websites, for which we collected 230,175 screenshots, including an estimated 115,000 ads. (For this study, we manually identified and labeled the ads for seven out of the 21 days.)

Finally, we present analysis, report findings, and share key insights. Section 4 summarizes some of our key findings. For additional analysis, discussion, and a full description of the methodology and data collection, we refer to the full version of this paper [8].

## 2 PROFILE CREATION

### 2.1 Persona design

**Characteristics.** Six different personas have been created for the purpose of imitating online user behavior. Every persona has been



*WPES '21, November 15, 2021, Virtual Event, Republic of Korea*



ACM ISBN 978-1-4503-8527-5/21/11…$15.00
https://doi.org/10.1145/3463676.3485617

Table 1: Summary of the six different personas.

| Tag | Gender | Age | Interest | Occupation | Civil Status | Kids |
|---|---|---|---|---|---|---|
| A | Female | 18 | Horses, celebrities, gardening | High School | Single | No |
| B | Female | 33 | Hair, fashion, DIY | Hairdresser | Single | No |
| C | Female | 51 | Movies, stock trading, interior design | Bank worker | Married | Yes |
| D | Male | 21 | American football, baseball, medicine | University | Relationship | No |
| E | Male | 37 | Cooking, traveling, electronics | History teacher | Divorced | Yes |
| F | Male | 68 | Birds, baking, crossword puzzles | Retired | Married | Yes |

assigned various characteristics: name, gender, age, three main interests, occupation, civil status, and parental status. Table 1 summarizes the key characteristics selected for each of the six personas.

**Search queries.** For each persona, we created 200 search phrases based on their individual key characteristics. 80% of the search queries were based on the interest categories and 20% were based on other personal information (e.g., whether they had children, were single, etc.). All personas use English as primary language.

**Stereotypes.** To simplify data collection and interpretation of the results, the personas were created to be fairly "stereotypical" in the sense that their interests tend to be commonly linked to their age and gender. These stereotypes are based on our own biases. We did not include personality traits such as religion, sexuality, or race. Finally, the personas were assigned common and widely applicable names: (A) Mary Johnson, (B) Jennifer Brown, (C) Patricia Jones, (D) James Davis, (E) John Anderson, (F) Robert Smith.

**Google accounts.** Six Google accounts were created to match the six personas. Several aspects of persona characteristics had to be revealed to Google when creating these accounts, including the name, birth date, and location. Only six out of 51 VMs utilized these profiles and hence also Google's log in function.

## 2.2 Framework design

We designed and implemented a Selenium-based framework that emulates the user behavior of a persona and collects a corresponding dataset. Figure 1 presents an overview of the framework. As input, the framework takes the search queries associated with one of the previously created personas. During a user session, the framework then considers one query at a time. The list of search phrases is shuffled every time the framework starts running.

**Weighted clickthrough.** After making a search query, the framework first filters the Google search result using a list of filter words. Thereafter, one of the available links is clicked, with the links being selected according to a Zipf-distributed probability distribution. This choice was motivated by previously observations for search queries and clickthrough rates [11, 17] and captures that the top results for a query see by far the most clicks.

**Post-search behavior.** After landing on a new webpage, there are four instances that can occur with different probabilities (*Px*).

- Backward-to-search (*P1*): User returns to the original Google results page.
- Forward-to-search (*P2*): User makes a new search query.
- Forward-to-browse (*P3*): User clicks link on the current page.
- Backward-to-browse (*P4*): User returns to the previous page.

Motivated by research by White and Drucker [30], we set these probabilities to *P1*=0.08, *P2*=0.21, *P3*=0.5 and *P4*=0.21.

**Think time.** We chose the think time to be 14 sec. with probability 0.45, 28 sec. with probability 0.35, and 56 sec. with probability 0.2. This choice results in a median similar to the 28.7 sec. observed by Ramakrishnan et al. [29].

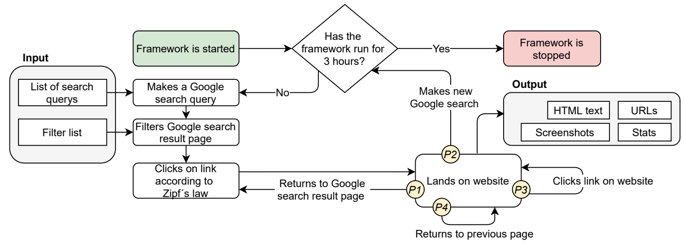


Figure 1: Overview of the framework design.

**Exception handling.** For simplicity, the framework accepts all pop-ups before continuing with additional tasks. If the framework runs into faulty pages (e.g., pages that do not load, HTML text that does not save, buttons that we cannot click) we return to the Google search page and continue with the next query. In the exceptional case that we run out of search phrases during a session, the framework repeats queries as necessary. This was very rare.

# 3 DATA COLLECTION

## 3.1 High-level experimental design

The experiments were split into two collection phases: the primary and secondary phase. The back-to-back phases lasted for 14+7 days. The data collection took place between 2021-04-24 to 2021-05-14.

**Primary phase.** We created a larger set of user profiles determined by the persona (personas A-F), the geographic location of the user (US East, US West, UK), what browser extension the user used (none, AdBlock [5], AdBlock Plus [26], Ghostery [2], CatBlock [1]), what browser the user used (Chrome, Firefox) and whether the user was logged in to their Google account or not. To ensure clean results, we run each profile inside a separate virtual machine (VM) given a unique public IP address. Appendix A (including Table 7) motivates and summarizes the setup used for each of the 51 VMs.

**Secondary phase.** At the start of the secondary phase, we "flipped" the extension-related configurations used by each user.

**System setup and collection details.** Using Microsoft Azure, we set up 51 VMs, one for each profile. The VMs had 2 vCPUs, 4 GiB memory, a standard SSD, and used Windows 10. The six logged in profiles remained logged in to their Google accounts for the duration of the data collection. In total, 33 VMs were placed in US East (Washington D.C.), 15 in UK South (London), and three in US West (San Jose). Each VM was adjusted to the corresponding time zone of their location. Furthermore, every session was launched at 04:00 PM (GMT+2) ± 2 hours, each day, and lasted for 3 hours. As a result of time zone differences, the daily data collection began at 09:00 AM ± 2 hours local time in US East, 06:00 AM ± 2 hours local time in US West, and 03:00 PM ± 2 hours local time in the UK.

## 3.2 Dataset creation and labeling

**Logged data.** During each session, log files were created containing information about the visited URLs, extracted HTML texts, and screenshots of the visited webpages. For each session, we also saved summary statistics about the number of clicked links and screenshots, for example, as well as information about potential exceptions. Finally, for the personas that were logged in to their Google account, we took manual screenshots of their Google Ad Personalization page at the end of each session.

**Table 2: Summary when the persona characteristics are added to the Google Ad Personalization page.**

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| **Interests** | Day 2 | Day 4 | Day 2 | Day 4 | Day 2 | Day 2 |
| **Occupation** | - | - | Day 2 | Day 2 | - | Day 6 |
| **Parental status** | Day 2 | Day 2 | Day 3 | Day 2 | - | - |
| **Relationship** | Day 8 | - | Day 2 | - | - | Day 2 |
| **Age** | Day 2 | Day 2 | Day 2 | Day 2 | Day 2 | Day 2 |
| **Gender** | Day 2 | Day 2 | Day 2 | Day 2 | Day 2 | Day 2 |
| **Match** | 83.3% | 66.7% | 100% | 83.3% | 50.0% | 83.3% |

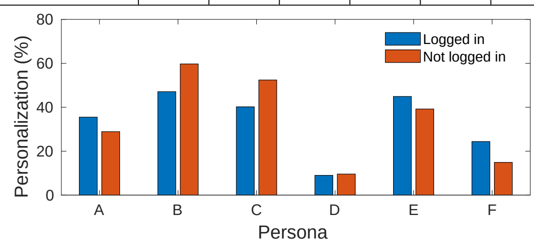


**Figure 2: Ad targeting based on whether logged in or not.**

**Identifying and categorizing ads.** To identify ads, the screenshots were manually evaluated. We next labeled each identified ad using the combined set of 18 interest categories of the six personas. Ads related to other persona traits (e.g., marital or parental status) were placed in the category "private life" of each persona. Ads that did not meet any of the above categories were labeled as "other".

## 3.3 Summary statistics

In total, we collected data for 1,071 sessions (51 VMs × 21 days), capturing 178,650 website visits and 203,175 screenshots. Given the manual effort required to annotate and label all screenshots, we focused on the screenshots from days 1, 6, 7, 13, 14, 15, and 21.

# 4 EXAMPLE RESULTS

## 4.1 Profile-based baseline comparison

**Google profile vs. persona profile.** Google managed to build a well-matching profile quickly. For example, at the latest, Google displays all persona interest categories by day 4 and most characteristics were identified within 3 days. Table 2 summarizes what day in the campaign that each type of information was learned.

**Logged in vs not.** As seen in Figure 2 the level of personalization is not significantly impacted by whether a user is logged in or not. For example, personas D and F consistently see low levels of targeting. The narrowness of these personas' interest categories is the most likely reason for this. It does not help that persona D is placed in the UK, as his sports interests (baseball and American football) are much less popular in the UK than in the US.

## 4.2 Detailed profile comparison

**Persona targeting.** Table 3 shows the fraction of ads observed by each persona (row in the table) that were associated with one of the – in some cases narrow – interest categories of the six personas (columns) as well as the "private life" category of each user.

There is a clear correlation of ads being related to the interest categories of each persona, resulting in stronger red color in the cells along the diagonal. However, we also observe several interest categories with very limited number of ads (mostly white columns). These categories may see limited over advertising budgets (e.g., "Birds" and "DIY") or companies may not target users in the region that user was located (e.g.,"American football" and "Baseball" ads in the UK). Yet, in all these cases the profile with the persona that best matched this interest saw the most ads for this category.

The level of personalization is also impacted by age and gender. For example, the age group 25-34 experiences most targeting (70%) and age group 65+ the least targeting (below 20%). Furthermore, the females (A, B, C) were exposed to twice the amount of targeted advertising seen by the males (D, E, F). Female personas are also exposed to more ads related to "fashion" than male personas. In contrast, male personas are exposed to almost twice the amount of "electronics" advertising than the female personas, and more than twice the amount of "traveling" advertising.

**Browser-based comparison.** While all profiles observed more ads per screenshot with Firefox than with Chrome (Table 4), the personalization experienced by the different profiles is relatively independent of whether we use Chrome or Firefox. For example, in both cases, personas D and F see the least personalization and persona B the most personalization. These results are summarized in Figure 3 and the findings are consistent also with what was seen when logged in to Chrome (see Figure 2). Also, the two attributes gender (Figure 4(a)) and age (Figure 4(b)) showed similarities in the level of personalization between the two browsers.

## 4.3 AdBlock and other extensions

**Extension comparisons.** While the use of extensions substantially reduced the number of ads seen by the users (on average by 87.1%), we have found that the fraction of personalized ads is higher than this baseline for three of the extensions. These results are summarized in Table 5 and show that the three most popular services are better at reducing more generic ads than personalized ads. Here, we also present $p$-values based on one-sided two-sample binomial hypothesis testing of the null hypothesis that the fractions of personalization seen by an extension is the same as without an extension. These results show statistical significance (with 95% confidence) for all extensions.

**Adding or removing extensions.** The secondary collection phase was implemented to glean some insights into the effects that adding or removing an extension may have on the personalization experienced by users. We observed (see full version [8]) that companies appear to successfully build user profiles also during the time a privacy enhancing extension is used. For example, a user inactivating an extension at the start of the secondary phase see more personalized ads on day 15 than what is observed by the corresponding user (without any extension running) on day 1. This shows that the tested extensions are not capable of fully stopping third-party tracking, despite some claiming so.

**Browser comparisons.** The relative level of personalization that users experience with the different extensions relative to when they do not use an extension is similar with Chrome and Firefox.

However, we did observe some differences in the effectiveness of the extensions when used with the two browsers. For example, with AdBlock, Chrome had a personalization level of 43.9% compared to 77.8% with Firefox (difference was significant at the 95% confidence; $p$=0.0099). For the other extensions, the differences were smaller and

Table 3: Fraction of ads associated with each of the interest categories of personas A-F. We also include a personal life category for each persona.

| | A | | | B | | | C | | | D | | | E | | | F | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Horses | Celebrities | Plants | Hair | Fashion | DIY | Movies | Stock-trading | Interior /Home | American Football | Baseball | Medicine | Cooking | Traveling | Electronics | Birds | Baking | Crosswords | Private life |
| **A** | 0.17 | 0.04 | 0.12 | 0.00 | 0.10 | 0.00 | 0.03 | 0.01 | 0.03 | 0.00 | 0.00 | 0.01 | 0.02 | 0.02 | 0.01 | 0.00 | 0.00 | 0.00 | 0.04 |
| **B** | 0.00 | 0.00 | 0.00 | 0.23 | 0.42 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.04 | 0.00 | 0.00 | 0.00 | 0.02 |
| **C** | 0.00 | 0.00 | 0.02 | 0.00 | 0.03 | 0.00 | 0.13 | 0.20 | 0.17 | 0.00 | 0.00 | 0.01 | 0.02 | 0.03 | 0.03 | 0.00 | 0.00 | 0.00 | 0.06 |
| **D** | 0.00 | 0.00 | 0.02 | 0.00 | 0.06 | 0.00 | 0.04 | 0.03 | 0.01 | 0.03 | 0.03 | 0.07 | 0.00 | 0.04 | 0.11 | 0.00 | 0.00 | 0.00 | 0.10 |
| **E** | 0.00 | 0.00 | 0.03 | 0.00 | 0.07 | 0.00 | 0.03 | 0.03 | 0.01 | 0.00 | 0.00 | 0.03 | 0.20 | 0.18 | 0.07 | 0.00 | 0.00 | 0.00 | 0.02 |
| **F** | 0.00 | 0.00 | 0.01 | 0.01 | 0.10 | 0.01 | 0.02 | 0.01 | 0.03 | 0.00 | 0.00 | 0.03 | 0.06 | 0.05 | 0.06 | 0.01 | 0.07 | 0.04 | 0.04 |

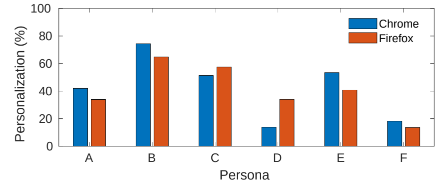


Figure 3: Browser: Chrome vs Firefox.

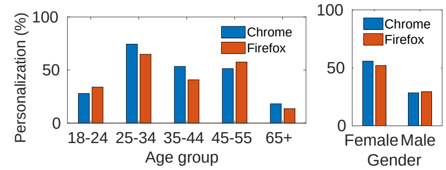


Figure 4: Age and gender.

Table 4: Ads/screenshot when using different browsers.

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| **Chrome** | 0.20 | 0.12 | 0.15 | 0.10 | 0.17 | 0.24 |
| **Firefox** | 0.28 | 0.16 | 0.32 | 0.11 | 0.21 | 0.26 |

Table 5: Personalized ads with an extension compared to without. An asterisk (*) is used to indicate that the result is statistically significant with 95% confidence (or higher).

| Extension | AB. | AB.+ | Gho. | Cat. | None |
|---|---|---|---|---|---|
| **Targeted** | 64.9% | 52.7% | 77.8% | 13.1% | 40.4% |
| **Ads/session** | 7.72 | 6.05 | 4.56 | 1.67 | 39.1 |
| **Ads/screenshot** | 0.050 | 0.038 | 0.026 | 0.011 | 0.219 |
| **p-value** | $<0.0001^*$ | $0.0078^*$ | $<0.0001^*$ | $0.0014^*$ | |

Table 6: Regional differences in the average number of ads observed and the level of personalization.

| Region | Ads/session | Ads/screenshot | Targeting |
|---|---|---|---|
| UK | 12.8 | 0.055 | 59.7% |
| US East | 38.8 | 0.143 | 46.0% |
| US West | 39.1 | 0.164 | 46.0% |

non-significant: 55.6% vs 49.8% (AdBlock Plus; $p$=0.2721), 70.2% vs 85.5% (Ghostery; $p$=0.0635, and 11.3% vs 14.8% (CatBlock; $p$=0.4067). Other than these smaller (mostly non-significant) differences, the results appear consistent across the browsers.

## 4.4 Regional differences

Table 6 summarizes our regional results. The UK users see much fewer ads than the corresponding US users. However, the level of personalization were much more similar, with UK-based users seeing only somewhat more personalization. We believe that the somewhat higher targeting when located in the UK may partially be situational. For example, perhaps some of the topics that are of interest to personas A, C, and E (e.g., horses) may be associated with bigger markets in the UK than in the US. When discussing these results, we note that we (like many users) leverage Google to find pages related to the users' interests. Since Google itself takes locality into account when directing users to different webpages, we expect that the observed differences may be contributed to by both differences in (1) the webpages visited and (2) the ads displayed when visiting a common set of webpages. We do not observe any significant differences between users in the eastern or western US.

**Campaigns.** During the study we observed several ad campaigns that were shown to a majority of the profiles, regardless of location and persona. Given the year that has been, it is perhaps not surprising that most of the identified campaigns centered around societal issues such as the coronavirus (e.g., "Corona: Wear a mask"). However, in general, most campaigns were US-focused and shown more to US-based users. Even the campaign most frequently seen by the UK users was for a US-based organization ("Feeding America").

# 5 RELATED WORK

**User tracking and advertisement.** Prior work has studied online privacy leakage [23, 24], third-party tracking [4, 16, 28], or the ability of a tracking service [14]. Both Carrascosa et al. [12] and Barford et al. [7] use alternative ways to build personas to evaluate the level of personalized advertisement. These works do not control and/or study many of the factors studied here.

**Extensions.** Many privacy and security risks associated with using the extensions have been identified [10, 13, 19, 20]. The browsers have mainly countered such risks through regulations [9, 22]. Others have studied the performance of adblockers [10, 18], the effectiveness of the lists they use [6], and geographic differences [25].

# 6 CONCLUSIONS

This paper presents a profile-based evaluation of the targeted advertising experienced by different users, including those that try to protect their integrity using privacy enhancing extensions. Using the emulation tool developed in the project, we performed a 21-day longitudinal measurement campaign, split into two phases. Using the datasets, we studied the personalization changes over time, starting from the day that a new profile is put online, how the level of personalization changes when adding or removing an extension, and how the results differ depending on the profile's persona (e.g., interest, occupation, age, gender, etc.), geographic location (US East, US West, UK), what browser extension they use (none, AdBlock, AdBlock Plus, Ghostery and CatBlock), what browser they use (Chrome, Firefox), and whether they are logged in to their Google account or not. For a complete analysis and discussion of the results, we refer to the full version of this paper [8].

**Acknowledgements:** This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

## REFERENCES

[1] [n.d.]. CatBlock. https://getcatblock.com/
[2] [n.d.]. Ghostery. https://www.ghostery.com/
[3] [n.d.]. Privacy Badger. https://privacybadger.org/
[4] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz. 2014. The Web Never Forgets: Persistent Tracking Mechanisms in the Wild. In *Proc. ACM CCS.*
[5] AdBlock. [n.d.]. About AdBlock. https://getadblock.com/
[6] M. Alrizah, S. Zhu, X. Xing, and G. Wang. 2019. Errors, misunderstandings, and attacks: Analyzing the crowdsourcing process of ad-blocking systems. In *Proc. IMC.* 230–244.
[7] P. Barford, I. Canadi, D. Krushevskaja, Q. Ma, and S. Muthukrishnan. 2014. Adscape: Harvesting and Analyzing Online Display Ads. In *Proc. WWW.*
[8] S. Bertmar, J. Gerhardsen, A. Ekblad, A. Höglund, J. Mineur, I. Öknegård Enavall, M-H. Le, and N. Carlsson. 2021. Who's Most Targeted and Does My New Adblocker Really Help: A Profile-based Evaluation of Personalized Advertising (extended). www.ida.liu.se/~nikca89/papers/wpes21.html
[9] A. Blondin. 9 Dec, 2020. Making Chrome extensions more private and secure. *Chrome - safety and security* (9 Dec, 2020).
[10] K. Borgolte and N. Feamster. 2020. Understanding the Performance Costs and Benefits of Privacy-focused Browser Extensions. In *Proc. WWW.*
[11] L. Breslau, P. Cao, Li Fan, G. Phillips, and S. Shenker. 1999. Web Caching and Zipf-Like Distributions: Evidence and Implications. In *Proc. IEEE INFOCOM.*
[12] J.M Carrascosa, J. Mikians, R.Cuevas, V. Erramilli, and N. Laoutaris. 2015. I Always Feel like Somebody's Watching Me: Measuring Online Behavioural Advertising. In *Proc. ACM CoNEXT.*
[13] Q. Chen and A. Kapravelos. 2018. Mystique: Uncovering Information Leakage from Browser Extensions. In *Proc. ACM CCS.*
[14] M. Degeling and J. Nierhoff. 2018. Tracking and Tricking a Profiler: Automated Measuring and Influencing of Bluekai's Interest Profiling. In *Proc. WPES.* 13.
[15] Disconnect. [n.d.]. No more trackers. https://disconnect.me/
[16] S. Englehardt and A. Narayanan. 2016. Online Tracking: A 1-Million-Site Measurement and Analysis. In *Proc. ACM CCS.* 1388–1401.
[17] M. J. Freedman. 2010. Experiences with CoralCDN: A Five-Year Operational View. In *Proc. NSDI.*
[18] K. Garimella, O. Kostakis, and M. Mathioudakis. 2017. Ad-blocking: A Study on Performance, Privacy and Counter-measures. In *Proc. ACM on Web Science Conference.* 259–262.
[19] C. Giuffrida, S. Ortolani, and B. Crispo. 2012. Memoirs of a browser: a cross-browser detection model for privacy-breaching extensions. In *Proc. ACM ASIACCS.*
[20] G. Gulyás, D. Somé, N. Bielova, and C. Castelluccia. 2018. To Extend or not to Extend: On the Uniqueness of Browser Extensions and Web Logins. In *Proc. WPES.*
[21] Raymond Hill. [n.d.]. uBlock Origin. https://chrome.google.com/webstore/detail/ublock-origin/cjpalhdlnbpafiamejdnhcphjbkeiagm?hl=en-GB/
[22] P. Kewisch, Rebloor, A. Wagner, A. Tasy, J. Villalobos, W. Bamberg, and Kmaglione. [n.d.]. Add-on Policies. https://extensionworkshop.com/documentation/publish/add-on-policies/
[23] A. Majumder and N. Shrivastava. 2013. Know Your Personalization: Learning Topic Level Personalization in Online Services. In *Proc. WWW.*
[24] D. Malandrino, A. Petta, V. Scarano, L. Serra, R. Spinelli, and B. Krishnamurthy. 2013. Privacy Awareness about Information Leakage: Who knows what about me?. In *Proc. WPES.*
[25] M. Malloy, M. McNamara, A. Cahn, and P. Barford. 2016. Ad blockers: Global prevalence and impact. In *Proc. IMC.*
[26] AdBlock Plus. [n.d.]. Surf the web with no annoying ads. https://adblockplus.org/
[27] E. Pujol, O. Hohlfeld, and A. Feldmann. 2015. Annoyed Users: Ads and Ad-Block Usage in the Wild. In *Proc. IMC.* 93–106.
[28] J. Purra and N. Carlsson. 2016. Third-party tracking on the web: A Swedish perspective. In *Proc. IEEE LCN.* IEEE, 28–34.
[29] R. Ramakrishnan, V. Shrawan, and P. Singh. 2017. Setting Realistic Think Times in Performance Testing - A Practitioner's Approach. In *Proc. ISEC.*
[30] R. White and S. Drucker. 2007. Investigating Behavioral Variability in Web Search. In *Proc. WWW.*


## A OVERVIEW OF VMs

**Primary phase setup:** We selected to consider one factor at a time, always including the three base profiles based on personas A, C, and E, which we located in eastern US, and that used Chrome without any extensions. First, to better understand differences between profiles we include results also for personas B, D, and E, which we placed (as a base case) in the UK. Second, we included this full set of profiles also when considering the impact of being logged in or not. Third and fourth, when studying the impact of location and the use of extension, we simply changed the location of the base profiles to the other two locations of considerations or applied one of the four extensions, respectively. Fifth, we repeated all tests comparing the impact of using different personas as well as all experiments with different extensions when using Firefox instead of Chrome. This was made possible since all four extensions worked for both Chrome and Firefox. Finally, we included three additional experiments with profiles B, D, and E. These were included to provide symmetric baseline results for each profile, when combining the results from the two phases of the data collection. Table 7 summarizes the set of VMs (with VMs number as per our internal numbering used in the datasets) and the secondary phase is described next.

**Secondary phase setup:** The profiles that used extensions during the primary collection were instead run with these uninstalled during the secondary phase. In contrast, the profiles that did not use any extension during the primary collection, added an extension for the secondary phase. Since we avoided having redundant VMs during the primary phase (which is the focus of most of our analysis!), we could not achieve the same level of coverage for each extension during the second phase. To allow some fair head-to-head comparisons also here, we therefore opted to tie the choice of extensions being added based on the persona used in each experiment. For VMs adding an extension for the secondary phase, we used the following assignments: AdBlock was added to VMs of personas A+D, AdBlock Plus to VMs of personas B+E, Ghostery to VMs of persona C, and CatBlock to VMs of persona F. Finally, to allow symmetry in the profile-based experiments looking at the impact of adding or removing extensions, we include VMs 37-39. These VMs ensure that we have a corresponding case where an extension is removed from a VM when there exists an experiment where that extension has been added to the same persona, location, and browser. (We opted to not design for the opposite to hold, since this would require redundant experiments during the primary phase, which is the primary focus of our study.)

**Table 7: The virtual machines with their set ups.**

| VM ID | Browser | Login | Persona | Region | Primary | Second |
|---|---|---|---|---|---|---|
| 1,3,5 | Chrome | No | A,C,E | East US | None | Ext. |
| 2,4,6 | Chrome | No | B,D,F | UK | None | Ext. |
| 7,9,11 | Chrome | Yes | A,C,E | East US | None | Ext. |
| 8,10,12 | Chrome | Yes | B,D,F | UK | None | Ext. |
| 19-21 | Chrome | No | A,C,E | UK | None | Ext. |
| 34-36 | Chrome | No | A,C,E | West US | None | Ext. |
| 22-24 | Chrome | No | A,C,E | East US | AdBlock | None |
| 25-27 | Chrome | No | A,C,E | East US | AdBlock Plus | None |
| 28-30 | Chrome | No | A,C,E | East US | Ghostery | None |
| 31-33 | Chrome | No | A,C,E | East US | CatBlock | None |
| 13,15,17 | Firefox | No | A,C,E | East US | None | Ext. |
| 14,16,18 | Firefox | No | B,D,E | UK | None | Ext. |
| 40-42 | Firefox | No | A, C, E | East US | AdBlock | None |
| 43-45 | Firefox | No | A, C, E | East US | AdBlock Plus | None |
| 46-48 | Firefox | No | A, C, E | East US | Ghostery | None |
| 49-51 | Firefox | No | A, C, E | East US | CatBlock | None |
| 37 | Chrome | No | B | UK | AdBlock Plus | None |
| 38 | Chrome | No | D | UK | AdBlock Plus | None |
| 39 | Chrome | No | F | UK | CatBlock | None |