sciendo

Proceedings on Privacy Enhancing Technologies ; 2020 (2):558–577

Gaston Pugliese*, Christian Riess, Freya Gassmann, and Zinaida Benenson

Long-Term Observation on Browser Fingerprinting: Users' Trackability and Perspective

Abstract: Browser fingerprinting as a tracking technique to recognize users based on their browsers' unique features or behavior has been known for more than a decade. We present the results of a 3-year online study on browser fingerprinting with more than 1,300 users. This is the first study with ground truth on user level, which allows the assessment of trackability based on fingerprints of multiple browsers and devices per user. Based on our longitudinal observations of 88,000 measurements with over 300 considered browser features, we optimized feature sets for mobile and desktop devices. Further, we conducted two user surveys to determine the representativeness of our user sample based on users' demographics and technical background, and to learn how users perceive browser fingerprinting and how they protect themselves.

Keywords: browser fingerprinting, tracking, privacy

DOI 10.2478/popets-2020-0041
Received 2019-08-31; revised 2019-12-15; accepted 2019-12-16.

1 Introduction

In 2020, the PETS paper *How Unique Is Your Web Browser?* [1] by Peter Eckersley will celebrate its 10th publication anniversary. It was the first paper describing a study on browser fingerprinting (PANOPTICLICK) that gained far-reaching attention and can thus be seen as the origin of this research field. Back then, Eckersley showed that a small set of eight browser characteristics, including the user-agent string and the list of installed plugins, were sufficient to uniquely recognize between 83.6% and 94.2% of browsers in his dataset.

Motivation. The collected browser fingerprints of PANOPTICLICK have not been published. Ten years later, after many related studies [2–6], research on browser fingerprinting still lacks available and appropriate data. Depending on the investigated aspects of browser fingerprinting, the requirements for data collection are quite extensive: long-term, large-scale, fine-grained, diversified, sometimes cross-browser and cross-device. These requirements apply to research studying the evolution of browser features over time, assessing users' trackability, or examining the effectiveness of countermeasures.

At the time of writing, we are aware of only two available datasets of browser fingerprints. Tillmann collected fingerprints in 2012 [2] and published a dataset that does not contain raw values of the most discriminating features, like fonts or plugins, and the data collection was performed over a short period of one month. Vastel et al. published an unfiltered sample of their data in 2018 of fingerprints collected via browser extensions [3]. Although the dataset contains fingerprints from 1.5 years, it lacks mobile devices, other browser families than Firefox and Chrome (after filtering) as well as precise timestamps and frequencies of observations.

Furthermore, most studies relied on cookies to recognize recurring browsers [1, 2, 4, 6]. This, however, is not a reliable way to establish ground truth for long-term observations, especially for privacy studies with allegedly savvy users: cookies can be easily deleted, either manually by the user or automatically by the browser, and they do not help to recognize users if they switch browsers or devices. Moreover, it diminishes the reliability of ground truth on fingerprints being *unique-by-entity* and *trackable* (Sec. 2.2).

In general, both the overall long-term trackability of users across multiple browsers and devices, and users' understanding of or actions against browser fingerprinting have received little attention so far. Furthermore, the representativeness of fingerprint datasets has only been investigated with respect to technical aspects [1–8], but not with respect to user demographics.

***Corresponding Author: Gaston Pugliese:** Friedrich-Alexander University Erlangen-Nürnberg, E-mail address: gaston.pugliese@cs.fau.de
Christian Riess: Friedrich-Alexander University Erlangen-Nürnberg, E-mail address: christian.riess@fau.de
Freya Gassmann: Saarland University, E-mail address: f.gassmann@mx.uni-saarland.de
Zinaida Benenson: Friedrich-Alexander University Erlangen-Nürnberg, E-mail address: zinaida.benenson@cs.fau.de

Finally, the impact of participation in the studies on users' perception of fingerprinting remains unknown.

Research Questions. Based on the considerations above, we present a 3-year online study with the goals to collect longitudinal fingerprint data and to investigate the users' perspective on browser fingerprinting for the first time. Our data collection (Sec. 3) establishes ground truth on user level instead of browser or device level. Thereby, we can link fingerprints to individual study participants over longer periods of time without relying on the persistence of client-side identifiers and regardless of the number of devices or browsers they used. Furthermore, we conducted two user surveys to determine the demographic representativeness of our user sample and to understand the users' perception of browser fingerprinting, the role of their study participation on this perception, and their protection measures. We aim to answer the following research questions:

- **RQ1:** How trackable are users based on their browser fingerprints regardless of the number of browsers and devices in use?
- **RQ2:** How do different feature sets perform regarding fingerprint stability and trackability of users?
- **RQ3:** Do demographics of users, as well as their technical background, privacy concerns and privacy behavior, correlate with their trackability?
- **RQ4:** How do users perceive browser fingerprinting, what is the role of study participation in this perception, and which countermeasures do they apply?

Hypotheses. For RQ3, we formulate the following six hypotheses, here combined into one sentence for brevity. **H1-6:** The following user characteristics are related to their trackability: (1) age, (2) gender, (3) education level, (4) computer science background, (5) privacy behavior, (6) privacy concerns.

Contributions. The main contributions of this paper cover technical findings as well as insights in users' perception of browser fingerprinting based on quantitative and qualitative analyses:

1. We present a novel long-term study on browser fingerprinting with ground truth on user level. Between 2016–2019, we collected 88,088 measurements of 305 browser features of 1,304 participants (Sec. 3).
2. We present two online surveys with study participants to assess their demographic characteristics and thus the representativeness of our sample as well as to study participants' privacy behavior, comprehension of browser fingerprinting, and applied countermeasures (Sec. 3.2, 5, and 6).

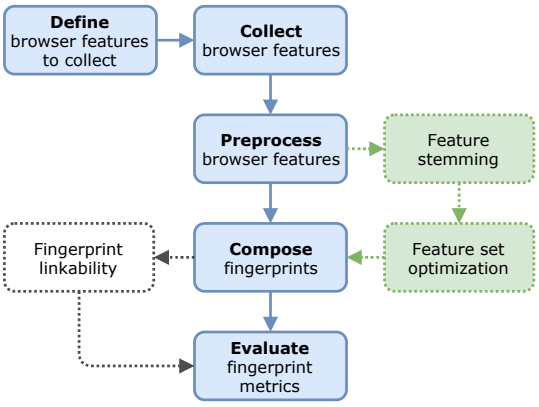

Fig. 1. Systemized workflow for browser fingerprinting

3. We present a simple, yet effective approach for optimizing feature sets towards different metrics (e.g., stability of fingerprints) for desktop and mobile devices. Further, we introduce *feature stemming* as a way to improve feature stability, e.g., by stripping off version substrings (Sec. 4).
4. We make a dataset of our long-term study available for research purposes (Sec. 7).

2 Background

A fingerprint is "a set of information elements that identifies a device or application instance" and fingerprinting is the "process of uniquely identifying" these entities [9]. For browser fingerprinting, these information elements (features) can be obtained *passively* from the client HTTP headers (e.g., user-agent string or language), and *actively* using a client-side script to collect information like screen resolution or plugins. Unlike cookies which are *stateful* identifiers stored on the client side, fingerprinting is considered a *stateless* tracking technique [10].

In the following, we review concepts and terminology of browser fingerprinting, and we provide an overview of studies that collected browser fingerprinting data since 2009 (Table 1).

2.1 Evaluating Browser Fingerprints

Figure 1 shows a workflow for browser fingerprinting: *(1)* The browser features that shall be collected are **defined** and the fingerprint script is implemented for the client and server side. *(2)* The fingerprint script is deployed to **collect** the clients' browser features. If ap-

plicable, these features are enriched with further stateful identifiers (e.g., cookies, session ID after authentication, personalized token in URL). Depending on the type of ground truth, a fingerprint can be linked to either an individual browser instance, device, or even user. *(3)* The collected browser features are **preprocessed** which may include *normalization* (e.g., screen resolution [8]) or *derivation* of additional information (e.g., user-agent parsing). Our work contributes to this step of the workflow and proposes *feature stemming* and *feature set optimization* to improve the stability of features and to compile an optimized feature set from collected features (Sec. 4). *(4)* The actual fingerprints are **composed** using a feature set. In practice, fingerprints can be handled as MD5 hashes, or as vectors of raw feature values. The latter enables further examinations like establishing *linkability* between evolved fingerprints [1, 3], which we do not consider in this work. *(5)* Finally, the collected fingerprints can be **evaluated** w.r.t. various metrics, e.g., the anonymity set size [11] or stability.

2.2 Formal Concepts

Considering browser fingerprinting formally, we denote the *feature set* consisting of n features as $\mathcal{X} = \{x_1, \ldots, x_n\}$ ($n \in \mathbb{N}$). Further, we denote the *feature value domain* of $x \in \mathcal{X}$, i.e., the set of all possible values of $x \in \mathcal{X}$ as $\mathcal{V}(x)$. As fingerprints are linked to entities (e.g., individual browsers, devices, or users) and they are observed at a point in time, we denote the *set of entities* as $\mathcal{E} = \{e_1, e_2, \ldots\}$, and the *time domain* as T.

We denote the *fingerprint type* of size k w.r.t. $\mathcal{X}$ as:

$$\mathcal{T} = (x_1, \ldots, x_k) \qquad (x_i \in \mathcal{X}, 1 \leq i \leq k) \ . \tag{1}$$

We denote the *fingerprint space* $\mathcal{V}$ of type $\mathcal{T}$ *w.r.t.* $\mathcal{X}$ as the cross product of $\mathcal{V}(x_i)$ ($x_i \in \mathcal{X}$):

$$\mathcal{V} = \mathcal{V}(x_1) \times \mathcal{V}(x_2) \times \ldots \mathcal{V}(x_k) \ . \tag{2}$$

The *set of all fingerprints w.r.t.* $\mathcal{T}$ is denoted as $\mathcal{F}$ and we define *fingerprints* ($f \in \mathcal{F}$) as follows:

$$f \colon \mathcal{E} \times T \to \mathcal{V} \ \text{ w.r.t. } \ \mathcal{T} \ . \tag{3}$$

Finally, for two timestamps t, t', and without loss of generality $t < t'$, we define the *stability period* $s(e,t)$ of fingerprint $f(e,t)$ as

$$\begin{aligned} s(e,t) &= \max(t' - t) \text{ s. t.} \\ \forall t'', t \leq t'' \leq t' &\colon f(e,t) = f(e,t'') = f(e,t') \ . \end{aligned} \tag{4}$$

Based on the previous equations, we derive following basic metrics for fingerprints.

Def. 1. *A fingerprint $f(e,t)$ w.r.t. $\mathcal{T}$ is **unique-by-entity** if, and only if, it is linked to a single entity, i.e.,*

$$\forall e' \in \mathcal{E}, \ \forall t' \in T, \ e' \neq e \ \Rightarrow \ f(e',t') \neq f(e,t) \ . \tag{5}$$

Def. 2. *A fingerprint $f(e,t)$ w.r.t. $\mathcal{T}$ is **unique-by-appearance** if, and only if, it was observed once, i.e.,*

$$\forall e' \in \mathcal{E}, \forall t' \in T, f(e',t') = f(e,t) \Rightarrow e' = e \wedge t' = t. \tag{6}$$

Def. 3. *A fingerprint $f(e,t)$ is **stable** if, and only if, its stability period is > 0.*

Def. 4. *A fingerprint $f(e,t)$ is **trackable** if, and only if, $f(e,t)$ is (i) unique-by-entity and (ii) stable.*

Analogously to Def. 4, an entity is considered *trackable*, if at least one trackable fingerprint is linked to it. In the sequel, whenever we use the terms such as "stable" or "trackable", we refer to the definitions above.

In related work, the concept of *uniqueness* is often vague as it is not stated whether feature values or fingerprints are *unique-by-appearance* or *unique-by-entity* w.r.t. to the corresponding ground truth.

2.3 Related Work

To our knowledge, Mayer [7] performed the first documented online study on browser fingerprinting. He collected 1,298 fingerprints from 1,328 browser instances and reported that 98.5% of the fingerprints were unique and thus 96.23% of the browsers uniquely identifiable.

Eckersley [1] presented PANOPTICLICK[1] and reported between 83.6%–94.2% of 470,161 fingerprints to be unique, depending on the availability of Flash or Java. Moreover, he established linkability between evolving fingerprints with an accuracy of 99.1%. Tillmann [2] collected 23,709 fingerprints from 18,692 browsers within one month, and reported that 92.57% of the fingerprints were unique.

Fifield and Egelman [6] uniquely identified 43% out of 1,016 browsers using the dimension of rendered font glyphs combined with the user-agent string.

Laperdrix et al. [4] collected 118,934 fingerprints on AMIUNIQUE[2] within three months, where 90% and 81% of the fingerprints from desktop and mobile devices were reported as unique, respectively.

1 https://panopticlick.eff.org/

2 https://amiunique.org/

Study	Year	Start	End	Features	Browsers	Users	Fingerprints	User Demographics	Dataset Availability	Ground truth: Cookie	Ground truth: IP Address	Ground truth: Browser ID	Ground truth: User ID
Mayer [7]	2009	02/2009	02/2009	3	1,328	-	1,298	○	○	●	○	○	○
Eckersley [1]	2010	01/2010	02/2010	8	-	-	470,161	○	○	●	●	○	○
Tillmann [2]	2013	11/2012	12/2012	48	18,692	-	23,709	○	●	●	○	○	○
Fifield and Egelman [6]	2015	-	-	43[1]	1,016	-	1,016	○	○	●	○	○	○
Laperdrix *et al.* [4]	2016	11/2014	02/2015	17	-	-	118,934	○	○	●	○	○	○
Cao *et al.* [8]	2017	-	-	49	-	1,903	3,615	○	○	●	○	○	●
Vastel *et al.* [3]	2018	07/2015	08/2017	17	1,905	-	98,598	○	●	○	○	●	○
Gómez-Boix *et al.* [5]	2018	12/2016	06/2017	17	-	-	2,067,942	○	○	●	○	○	○
This study	2019	02/2016	02/2019	305[2]	-	1,304	88,088	●	●	○	○	○	●

[1] Width and height of 43 Unicode code points, each rendered in six default CSS font families; [2] including parsed and derived ones

Table 1. Overview on browser fingerprinting studies that collected datasets of fingerprints between 2009 and 2019 including this study; indicating the publication year, start and end date of data collection as well as the number of considered features and fingerprints after data cleansing and filtering, the number of distinct browsers or users observed (if explicitly specified), whether user demographics were collected, the public availability of the collected data, and the ground truth for each dataset (● yes, ○ no)

Cao et al. [8] investigated cross-browser fingerprinting and collected 3,615 fingerprints from 1,903 users. They proposed novel rendering tasks for canvas fingerprinting and reported an identification rate of 99.24% on unique fingerprints.

Vastel et al. [3] used browser extensions for Firefox and Chrome to collect fingerprints, which is a major improvement towards establishing ground truth for long-term observations. They proposed FP-STALKER, an algorithm to establish linkability between fingerprints despite changing features.

Gómez-Boix et al. [5] collected 2,067,942 fingerprints from a real-world population on a popular french website using the same feature set as [3, 12]. They reported only 33.6% and 18.5% of the fingerprints from desktop and mobile devices to be unique. This notable lower share of unique fingerprints indicates the need for real-world samples of browser fingerprints for a thorough assessment of the threat of browser fingerprinting on individuals' privacy.

Compared to our work, most studies above collected the data only within six months or less, except [3]. Our evaluations are based on data collected within three years, which we believe can provide new insights on users' long-term trackability. In contrast to most prior works, except for [8, 13], we did not rely on cookies to recognize recurring users. We used personalized links to establish ground truth on user level, like [8], and allowed users to choose the browsers *and* devices they want to use freely, unlike [8]. To our knowledge, our study is the first to have ground truth on user level while allowing multiple devices per user.

Previous studies did not collect demographics of their participants. Yet, we assume that demographics are important to assess the representativeness of a study: Studies with explicitly recruited users report a much higher share of unique fingerprints than those with real-world users [5]. Demographic characteristics may be one reason for this discrepancy. Recruited user samples are likely to have a higher share of students and professionals with technical background as they are more likely to be interested in browser fingerprinting. To our knowledge, our study is the first to provide such data.

Furthermore, our study is the first to conduct user surveys to investigate how users perceive browser fingerprinting, how they protect themselves against it, and how their participation in the study impacts their view on browser fingerprinting and privacy.

Finally, while prior studies considered between 3 and 49 browser features (see Table 1), our study considers 305 features, including derived ones (e.g., via user-agent parsing or feature stemming). Unlike other studies, we do not use a predefined feature set, but instead investigate how different feature sets perform regarding fingerprints' stability and how they affect users' trackability on the long run.

3 Method

Below we describe our study design: How we collected browser fingerprints, how we conducted user surveys, and how we compiled the final dataset for evaluation.

3.1 Study Design

We designed an online user study where participants register themselves on our website: https://browser-fingerprint.cs.fau.de/. In the following, we describe the main components and considerations during the design of our study.

3.1.1 Establishing Ground Truth

Participation in a study on browser fingerprinting might encourage users to experiment with their browsers to become undistinguishable, or to use multiple browsers and devices out of curiosity. As we aimed for a long-term observation of fingerprints from multiple devices and browsers per user, we required a level of ground truth that is more reliable than that provided using cookies. We refrained from user accounts, apps, or browser extensions to establish ground truth as it would have raised the hurdle for participation (e.g., remembering credentials), require an additional communication channel to remind the users to submit measurements, or limit our sample to specific devices or browsers [8, 14].

We decided to require users to sign up with their email addresses. Afterwards, they verified their email addresses using a verification link sent to them. Thereby, we establish ground truth on user level, enable users to use as many browsers and devices as they wish, and also control the reminders for periodic measurements by sending weekly emails with personalized links to them.

3.1.2 Ethical Considerations

The study received an approval from the data protection office at our institution. We attached great importance to transparency towards participants about browser fingerprinting, the purpose of our study, and the data we collect. We fully disclosed our study method on our website, including all conceptional and technical details (e.g., feature collection, user surveys, data storage).

During the entire study, participants were free to provide their fingerprints by visiting the weekly links we sent them via email. At any given point in time, participants were able to quit their participation by visiting an unsubscription link we added to each of the weekly emails. If participants decided to quit, we automatically removed their email addresses from our database. All data is stored on a server hosted at our institution, with only project researchers having access to the data.

3.1.3 Fingerprinting and User Experience

After participants verified their email address, they received a new personalized link once a week via email which referred to a subpage of our study website where their browser features were collected. Each weekly link had a validity period of one week to reduce the potential effect of publicly shared links that could distort the evaluation of individual participants retrospectively.

One of our goals was to investigate how users perceive browser fingerprinting, and whether our study has impact on this perception. Hence, we provided information about browser fingerprinting and our study as well as FAQs written in non-technical terms. The website, surveys, and email texts were in English and German.

After collecting the browser features using a custom script, we informed the participants about the uniqueness and recognizability of their fingerprints and provided an overview on all of their features. We further provided descriptive statistics about our study.

To provide another incentive for participation on a regular basis, every four weeks we sent an individual study report to each participant. From these reports, the participants could learn for how long their three most stable fingerprints were trackable, for desktop and mobile devices separately, and how their trackability compares to the results of other participants.

3.1.4 User Survey 1: Demographics and Background

To determine the representativeness of our user sample, and to answer the research question RQ3 about the relation between various user characteristics and their trackability, we designed a user study to be administered shortly after the users registered for the study. The survey was tested with seven experts on usable security and five private contacts with non-technical background, and iteratively improved during the tests.

We collected year of birth, gender, country of residence, current occupation, and the highest level of education. We also asked whether they study or work in computer science or a related discipline, and whether they have ever heard of browser fingerprinting before.

We measured *privacy concerns* using the Westin index [15] and asked the following questions about the *privacy behavior* on the Web:

- Have you ever used following privacy measures: Do Not Track, Tor (e.g., Tor browser), private mode in browsers (privacy or incognito mode), deleting cookies, denying third-party cookies?
- Have you ever used following browser extensions: Ad Blocker, NoScript, BetterPrivacy, HTTPS Everywhere, Privacy Badger, Ghostery, Disconnect, uBlock Origin?

For each privacy behavior, users could answer with "yes", "no" or "don't know". We computed a *privacy behavior index* based on the number of "yes" answers.

3.1.5 User Survey 2: Impact of Our Study

To answer RQ4 about users' perception of browser fingerprinting, applied countermeasures, and impact of our study, we developed a second user study. The survey was tested with twelve experts on usable security and five laypersons, and iteratively improved during the tests.

Understanding, Concern, Protection. We gathered quantitative data on users' understanding and perception of browser fingerprinting by asking them to indicate their agreement or disagreement with the following statements on a 5-point Likert scale from *strongly disagree* to *strongly agree*: *(1)* Browser fingerprinting can be used by websites to recognize me. *(2)* I think that I understand how browser fingerprinting works. *(3)* I am concerned that websites and companies try to fingerprint my browser. *(4)* I think most websites on the Web use browser fingerprinting. *(5)* I think websites that I frequently visit use browser fingerprinting. *(6)* It is important for me to be protected from browser fingerprinting. *(7)* I am capable of protecting myself from browser fingerprinting. *(8)* Protecting myself from browser fingerprinting requires a lot of effort.

The items were presented to each user in a randomized order to counter the influence of the previous items on the answers to the subsequent ones. In the above order, items 1 and 2 measure users' (perceived) understanding of fingerprinting, items 3-5 refer to the concern about fingerprinting and perception of its spread, and items 6-8 ask users' opinion about protection measures.

Countermeasures. To identify applied protection measures, we asked the following question, where participants could enter their countermeasures into predefined boxes (number of boxes was not fixed): Did you try to protect yourself from browser fingerprinting? If yes, what did you try (at least occasionally)? Please specify as many measures as you can think of.

Study impact. We asked the following questions, where the users first indicated "yes" or "no", and then could explain their reasons (free text): *(1)* Did you experience or learn something new by participating in our study? *(2)* Did your study participation change any of your thoughts/feelings regarding the Web? *(3)* Did your study participation change your behavior on the Web?

3.2 Data Collection

Recruiting and Initial Data Sample. Our goal was to recruit as many participants as possible from the general public. Thus, we used mailing lists of international universities, press releases, and private contacts of researchers. Moreover, participants could share our study via email or via privacy-respecting Twitter, Facebook and Google+ buttons with a predefined text. We also used announcements at computer science and security conferences, mailing lists, and forums.

Within 1,111 days between February 9, 2016 and February 23, 2019, we collected a total of 124,046 measurements from 2,315 participants. Each measurement consists of 305 features (App. D). We do not call the data *fingerprints* yet, but *measurements*, because the number of distinct fingerprints is depended on the feature set used to compose the fingerprints (Sec. 2, 4).

Data Cleansing. As we aim to assess users' trackability on the long run, we had to filter our raw dataset to remove the *noise* induced by exploitation of our study design and by users' curiosity. Using a keyed-hash stored additionally for each email address, we merged 42 participants that re-subscribed with the same email address after unsubscribing. We discarded 2,421 measurements resulted from testing and debugging with 11 of our own subscriptions to the study. We removed 11,294 first-week measurements of all participants to reduce the effect of their initial curiosity and experimentation during the first week on later evaluations. This resulted in 266 of the participants having no measurements left. To ensure a minimum observation time of at least four weeks with four measurements, we had to remove 676 participants (2,173 measurements). Further, we removed 19,647 measurements from the upper 1% of partici-

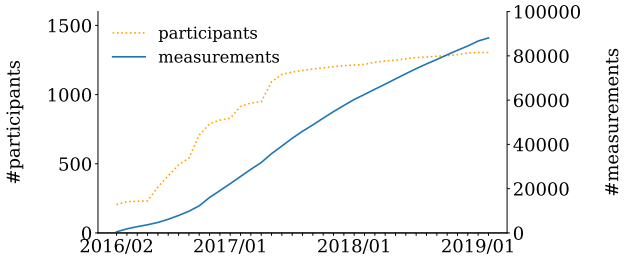

Fig. 2. Growth of filtered dataset during course of study

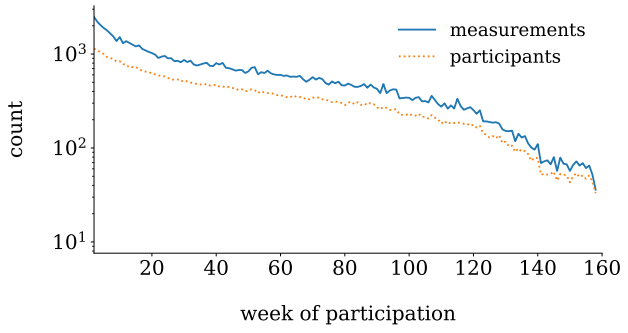

Fig. 3. Collected data per participation week in filtered dataset

		n	%
Browser	Delete cookies	1,200	94.1
	Use of private Mode	1,080	84.7
	Denying third-party cookies	1,036	81.3
	Set *Do-not-track* flag	916	71.8
	Tor / Tor browser	605	47.5
Extensions	AdBlock	1,058	83.0
	NoScript	666	52.2
	Ghostery	526	41.3
	HTTPS Everywhere	479	37.6
	uBlock	340	26.7
	Privacy Badger	238	18.7
	Better Privacy	234	18.4
	Disconnect	195	15.3

Table 2. Participants' *privacy behavior* (N=1,275)

pants with the most measurements (16) as their amount and frequency of measurements appeared dubious (e.g., due to sharing of personalized links, or automated requests). Finally, we discarded 423 measurements from clients that were neither desktop nor mobile devices (e.g., crawler, command-line tools).

Dataset for Evaluation. Our final dataset consists of 88,088 measurements from 1,304 participants and is used in the sequel. On average, participants provided 67.6 measurements (σ=78.4) over a period of 63.2 weeks (σ=47.7). Figure 2 shows the growth of the number of participants and measurements in our filtered dataset, and Figure 3 shows the total number of participants and measurements per participation week.

Sample Characteristics. In total, 1,275 study participants (97.8%) answered the first user survey. 79.1% of participants were from Germany (1,008); the remaining shares within the top 5 countries were: 3.8% U.S. (48), 2.2% Netherlands (28), 1.6% UK (20), and 1.0% Switzerland (13). The majority of participants were male (76.5%), between 30 and 49 years old (40.2%), have an academic background (64.8%), are employed (47.7%), study or work in computer science or a related field (57.5%), and already knew browser fingerprinting (68.5%). A full overview of our participants' demographics is shown in Appendix A.

Based on the Westin index [15], 60.9% of the participants were categorized as *privacy fundamentalist*, 36.5% as *pragmatic*, and 2.6% as *unaware*. Regarding users' *privacy behavior*, Table 2 shows the results on whether the 13 privacy measures we asked for have ever been used by the participants. Counting the "yes" answers, the mean *privacy behavior index* is 6.7.

4 Trackability of Participants

In this section, we present a data-driven approach to select suitable tracking features. We first create stabilized versions of features, and subsequently perform an automated feature selection. The feature selection greedily maximizes an objective function that can be adjusted to the target application and for different device types. We propose objective functions to maximize the number of trackable users, and to maximize feature stability.

Our results outperform the hand-crafted feature sets of Panopticlick [1] and AmIUnique [4]. Moreover, we believe that automated feature selection can provide a more realistic view on how the trackability of users might be improved by privacy-violating websites even without using *linkability* techniques [1, 3].

4.1 Feature Stemming

Several browser features change regularly. This involves, e.g., changing version strings after system updates, or varying screen resolution upon connecting the device to an external monitor. Vastel *et al.* distinguish changes like version strings as *automatic evolutions*, changes like varying screen resolution as *context-dependent evolu-*

tions, and changes like disabling cookies or enabling "Do-not-track" as *user-triggered evolutions* [3].

Particularly automatic evolution can negatively impact the stability of browser fingerprints. Hence, we manually selected features and removed their variable elements (such as version strings via regular expressions). We denote this normalization as *feature stemming* in reference to linguistic processing to reduce "words with the same root [...] to a common form [...] by stripping each word of its derivational and inflectional suffixes" [16]. Appendix B shows examples for the erasure of version substrings, assignment of IDs to plugins and MIME types, and sorting of list-like features in alphabetical order to gain robustness against targeted order randomization.

As shown in Appendix C, *stemmed* versions of features have been selected by the data-driven feature selection presented in the next section.

4.2 Feature Set Optimization

We propose two objective functions for feature set optimization, namely the number of trackable users and the average stability period. Both objective functions might be suitable choices for real-world tracking tasks. Thus, we consider them equally suitable for privacy assessments of users' trackability.

Number of Trackable Users (J_u). To maximize the number of trackable users, we define a mapping function $\phi : \mathcal{F} \mapsto \mathcal{R}$ that reduces the set of all fingerprints $\mathcal{F}$ to the set of trackable fingerprints $\mathcal{R}$. The resulting set $\mathcal{R}$ depends on a specific choice of a fingerprint type $\mathcal{T}$. Therefore, if we seek to maximize the number of trackable users, the goal is to find via the objective function J_u an optimum fingerprint type $\mathcal{T}^*$,

$$J_\mathrm{u} : \mathcal{T}^* = \arg\max_{\mathcal{T}} |\psi(\phi(\mathcal{T}))| \ , \tag{7}$$

where the mapping $\psi : (e,t) \mapsto (e)$, $(e,t) \in \mathcal{R}$ extracts the entity from a trackable fingerprint.

Stability of Trackable Fingerprints (J_s). To maximize the stability period of trackable fingerprints (Def. 4), we propose the objective function J_s to seek

$$J_\mathrm{s} : \mathcal{T}^* = \arg\max_{\mathcal{T}} \frac{\sum s(\phi(\mathcal{T}))}{|\mathcal{E}|} \ , \tag{8}$$

where the function $s : (e,t) \mapsto (t_\mathrm{max})$ calculates the stability period of a fingerprint as defined in Eqn. 4. Note that $|\mathcal{E}|$ is constant across fingerprint types $\mathcal{T}$ in Eqn. 8, and can hence be omitted during optimization.

Greedy Sequential Search. In practice, using the objective functions Eqn. 7 or Eqn. 8 in a brute-force search through all subsets of features is infeasible. However, we can perform a search for a sufficiently good type. This is referred to as *feature selection* in the pattern recognition literature, with various existing approaches that vary in computational effort and reliability of the results [17]. We are not aware of an efficient search strategy that guarantees convergence to the *global* maximum of Eqn. 7 or Eqn. 8. Hence, we resort to a greedy sequential search. A greedy search is only guaranteed to converge to a local optimum. Yet, the presented empirical results outperform manual feature selection strategies, which illustrates the benefit of a data-driven optimization over hand-crafted solutions.

In detail, the search iteratively enlarges the cardinality of the fingerprint type by adding that feature that yields the largest improvement to the objective function. Thus, it starts with an empty fingerprint type $\mathcal{T}^{(0)}$, $|\mathcal{T}^{(0)}| = 0$. In iteration i, the type expansion is

$$\mathcal{T}^{(i)} = \mathcal{T}^{(i-1)} \cup x_\mathrm{opt} \ , \text{where} \tag{9}$$

$$x_\mathrm{opt} \notin \mathcal{T}^{(i-1)}, \tag{10}$$

$$x_\mathrm{opt} = \arg\max_{x} \ J \quad \text{for } J \in \{J_\mathrm{u}, J_\mathrm{s}\} \ , \tag{11}$$

and J_u or J_s chosen as in Eqn. 7 and Eqn. 8, respectively. The iteration is terminated when there is no feature found that further improves the chosen objective function. If multiple features offer identical improvement in one iteration, we choose the lexicographically first.

4.2.1 Evaluation on Our Dataset

The dataset is split in non-overlapping training and test sets by odd and even user IDs with 652 users in both sets. The splits contain 45,063 and 43,025 measurements over a period of 157.6 and 157.3 weeks, respectively. The test set measurements consist of 29,989 desktop and 13,036 mobile measurements of 621 and 432 users. Both sets are statistically similar, with shares of mobile devices 29.3% and 30.3%, JavaScript enabled 86% and 85.7%, Flash enabled 13.7% and 13%, and on average 69 and 66 measurements per participant, respectively.

The training set is used to select *device-dependent* (desktop $\mathcal{T}^D$, mobile $\mathcal{T}^M$) and *device-independent* (total $\mathcal{T}^T$) feature sets. For each device, we select one feature set to maximize the number of trackable participants using Eqn. 7 (denoted as $\mathcal{T}_\mathrm{u}^D$, $\mathcal{T}_\mathrm{u}^M$, $\mathcal{T}_\mathrm{u}^T$), and one feature set to maximize stability using Eqn. 8 (denoted as $\mathcal{T}_\mathrm{s}^D$, $\mathcal{T}_\mathrm{s}^M$, $\mathcal{T}_\mathrm{s}^T$). Also on the test set, we compare to the

	Device type	Desktop				Mobile				Total			
	Feature set	[1]	[4]	$\mathcal{T}_{\mathrm{u}}^{D}$	$\mathcal{T}_{\mathrm{s}}^{D}$	[1]	[4]	$\mathcal{T}_{\mathrm{u}}^{M}$	$\mathcal{T}_{\mathrm{s}}^{M}$	[1]	[4]	$\mathcal{T}_{\mathrm{u}}^{T}$	$\mathcal{T}_{\mathrm{s}}^{T}$
User (%)	w/ unique FPs (appear.)	95.7	97.1	87.3	91.0	80.3	88	88.7	77.1	95.7	97.4	96.8	91.4
	w/ unique FPs (entity)	98.1	**99.4**	98.6	98.7	89.1	**94.4**	94.2	88.4	98.2	**99.5**	99.4	98.3
	w/ trackable FPs	84.4	85.7	**89.2**	**89.2**	67.6	**72.9**	72.7	64.6	91.3	93.1	**94.5**	**94.5**
Stability trackable FPs (weeks)	Mean of means p. user	3.1	3.4	9.6	**11.6**	3.3	3.1	3.1	**10.7**	3.2	3.3	3.7	**11.9**
	Std. of means p. user	3.3	3.9	9.3	10.5	6.4	5.1	5.1	11.8	4.1	3.8	4.6	10.8
	$q25$ of means p. user	1.3	1.5	4.1	**4.5**	1.0	1.1	1.1	**2.1**	1.4	1.5	1.8	**4.6**
	$q50$ of means p. user	2.2	2.4	7.9	**9.0**	2.2	2.3	2.3	**7.0**	2.3	2.4	2.9	**9.4**
	$q75$ of means p. user	3.8	3.9	12.2	**16.0**	3.5	3.3	3.4	**14.4**	3.8	3.8	4.2	**15.7**
	Mean of maxima p. user	8.2	8.6	21.1	**25.7**	6.8	6.8	6.8	**20.2**	9.2	9.5	10.0	**27.2**
	Std. of max. p. user	12.3	12.7	19.5	23.7	12.1	10.9	10.9	24.2	13.9	13.8	14.3	25.0
	$q25$ of max. p. user	2.1	2.4	6.0	**7.0**	1.9	2.0	2.0	**3.0**	2.4	2.9	3.2	**8.0**
	$q50$ of max. p. user	4.8	5.0	15.2	**18.1**	4.0	4.1	4.2	**10.0**	5.0	5.5	6.0	**19.1**
	$q75$ of max. p. user	8.1	8.3	29.3	**40.0**	7.0	8.0	8.0	**27.8**	9.0	9.0	10.0	**42.0**
FPs (%)	Distinct FPs (n)	12,330	13,801	7,473	8,029	3,935	4,793	4,825	4,308	16,265	18,594	16,541	9,822
	Unique FPs (appear.)	57.0	60.7	49.2	53.2	47.9	50.9	51.4	70.1	54.8	58.2	53.9	49.1
	Unique FPs (entity)	94.4	97.7	97.4	**97.8**	90.1	91.8	91.6	**94.1**	93.4	**96.2**	95.8	95.4
	Trackable FPs	37.1	36.7	**48.0**	44.4	**41.8**	40.6	39.8	23.7	38.2	37.7	41.6	**46.1**

Table 3. Feature set optimization towards "*users with trackable fingerprints*" (J_{u}) and "*stability of trackable fingerprints*" (J_{s}). Feature sets of *Panopticlick* [1], *AmIUnique* [4], and *computed feature sets* ($\mathcal{T}^{D}$, $\mathcal{T}^{M}$, $\mathcal{T}^{T}$) applied on the test split of our dataset. **Bold** values indicate best result(s) p. row. Highlighted rows indicate the optimization criterion for $\mathcal{T}^{D}$, $\mathcal{T}^{M}$, $\mathcal{T}^{T}$ *w.r.t.* J_{u} and J_{s}.

prior works with preset feature sets Panopticlick [1] and AmIUnique [4]. For the latter, the feature on the order of HTTP headers is omitted, as our data does not provide this information (discussed in Section 7).

Results. Our feature selection chose from the 305 features with J_{u} and J_{s} on desktop data 9 and 15 features, on mobile data 9 and 15 features, and on the full dataset 11 and 24 features (Appendix C).

The results of our feature selection are shown in Table 3. Three results are shown in vertical order: results on the number of tracked participants, results on the stability of the fingerprints, and statistics on the number of fingerprints. We discuss these results below.

The first gray row is the central result on the number of tracked users. Here, both proposed feature selection methods achieve with 89.2% by some margin the best result for desktop devices. On mobile devices, AmIUnique is with 72.9% slightly better than our feature set $\mathcal{T}_{\mathrm{u}}^{M}$ with 72.7%. On the whole dataset ("Total"), both proposed methods achieve the best results (94.5%).

The second gray row is the central result on the stability of fingerprints. As a performance metric, we calculate the mean over the average stability periods of each user. When optimizing for the number of tracked users via J_{u}, the stability is roughly comparable to Panopticlick and AmIUnique. However, this changes completely when optimizing for the stability via J_{s}: here, the achieved stabilities range between 10.7 weeks for mobile devices and 11.9 weeks for the whole dataset, which outperforms the related methods by a factor of about 3. The reported standard deviations are also considerably larger, since the fingerprints of some users are remarkably stable fingerprints. This is further illustrated in the quartiles in the lines below. For the mean over the maximally stable fingerprints per user, the third quartile of 42 weeks on the whole dataset shows that some users have extremely stable fingerprints.

The fingerprint statistics in the third part of Table 3 show that our method extracts on a considerably lower number of distinct fingerprints on desktop devices, and on mobile devices a slightly larger number. High percentages in the following row "unique-by-appearance" indicate that a larger number of fingerprints is useless for tracking as they only appear once and thus have no stability. By tendency higher percentages here coincide with lower number of tracked users and reduced stability. Conversely, a larger percentage in "unique-by-entity" and "trackable" fingerprints improves the tracking result. Thus, it is interesting to note that the proposed feature selection apparently runs into a local optimum when working with mobile device data, by producing a large share of unique-by-appearance fingerprints and as a consequence a reduced percentage of trackable fingerprints. On the other hand, the feature selection

works remarkably well on desktop data and the whole dataset with a somewhat lower number of distinct fingerprints, but these fingerprints are more precise with respect to the number of tracked users and fingerprint stability.

4.2.2 Evaluation on the FP-STALKER Dataset

We also evaluate our feature selection on the FP-STALKER dataset[3] which is the only available large fingerprint dataset at the time of writing (see Table 1). As shown in Figure 1, linkability should be applied after the preprocessing (which includes feature stemming and feature set optimization) and fingerprint composition, and is thus not considered in this work. However, the FP-STALKER dataset is suitable for the evaluation of feature set optimization, as we explain below.

Dataset. The fingerprints of 1,819 browser instances were collected using browser extensions for Chrome and Firefox. We discard all other browser instances, and also browsers with more than a single browser and operating system family, as this is a strong indicator for spoofing. We further remove fingerprints that are inconsistent according to FP-STALKER, that are unique-by-appearance, that have tiny screen resolutions (e.g., 8x8), and the upper 1% of browsers with more than four canvas fingerprints. This left 1,198 browsers and 4,816 distinct fingerprints for evaluation.

Since the dataset is pre-grouped into fingerprints using a given feature set, we use the timestamps to split the fingerprints into individual measurements using each fingerprint's time of first appearance, update, and last appearance, resulting in 14,887 measurements.

We split the data into training and test, each with 599 browser instances with 7,025 and 7,862 measurements from within 72.7 weeks, respectively. We also derive features via user-agent parsing, screen resolution normalization [8], and feature stemming, and split the *WebGL fingerprint* into individual features.

Results. As shown in Table 4, optimizing towards J_u provides almost identical stability as the initial FP-STALKER feature set (baseline). In both cases, the average stability of trackable fingerprints per browser is 1.8 weeks. In the upper quartile of average and maxima stabilities, J_u outperforms FP-STALKER by 0.1 weeks, where one browser less is trackable for J_u. The share of trackable fingerprints is equal to the baseline (95%).

	Criterion	Baseline	J_u	J_s
	Features (n)	18	8	13
Brow. (%)	w/ unique FPs (appear.)	16.5	16.4	11.5
	w/ unique FPs (entity)	**97.8**	97.5	95.7
	w/ trackable FPs	**91.0**	90.8	89.1
Stability trackable FPs (weeks)	Mean of means p. browser	1.8	1.8	**3.7**
	Std. of means p. browser	4.1	4.1	5.6
	q25 of means p. browser	**0.2**	**0.2**	**0.2**
	q50 of means p. browser	0.6	0.5	**1.5**
	q75 of means p. browser	1.4	1.5	**4.4**
	Mean of maxima p. browser	2.9	2.9	**6.3**
	Std. of maxima p. browser	5.1	5.0	8.3
	q25 of maxima p. browser	0.2	0.2	**0.3**
	q50 of maxima p. browser	1.3	1.3	**2.6**
	q75 of maxima p. browser	3.1	3.2	**9.4**
FPs (%)	Distinct FPs (n)	2,513	2,501	1,648
	Unique FPs (appear.)	4.6	4.6	4.6
	Unique FPs (entity)	**99.6**	**99.6**	98.5
	Trackable FPs	**95.0**	**95.0**	93.9

Table 4. Feature set optimization on FP-STALKER dataset compared to its initial feature set (baseline) [3]. Two feature sets optimized towards criteria "*entities with trackable fingerprints*" (J_u) and "*stability of trackable fingerprints*" (J_s) were computed on training set and applied on the test split. **Bold** values indicate best result(s) per row. Highlighted rows indicate the optimization criteria.

J_u composes 2,501 distinct fingerprints using 8 features, while the baseline yields 2,513 using 18 features.

Optimizing directly towards the stability of trackable fingerprints (J_s) yields an average stability of 3.7 weeks, hence outperforms the baseline by a factor of 2. On $q75$, J_s achieves 4.4 weeks on the average stabilities and 9.4 weeks on the maximum stabilities, which is 3 times higher than the baseline. The number of distinct fingerprints of the baseline is 2,513, which is considerably larger than the 1,648 distinct fingerprints that J_s composes out of 7,862 measurements on the test set. The trade-off for this improvement in stability is that 11 browsers are less trackable with J_s, and the share of trackable fingerprints is slightly lower (93.9% vs 95%).

The results show that a data-driven feature selection can considerably improve a given figure of merit, in this case stability.

5 Trackability Factors

We conducted quantile regressions [18] to test Hypotheses H1-6 (Sec. 1) about the relations between user characteristics and trackability.

3 https://github.com/Spirals-Team/FPStalker

	Model 1			Model 2			Model 3		
	q25	q50	q75	q25	q50	q75	q25	q50	q75
Age	0.00964^{+}	0.0147***	0.0245***	0.0130*	0.0157*	0.0171^{+}	0.0128*	0.0189**	0.0272***
	(1.93)	*(3.72)*	*(3.55)*	*(2.19)*	*(2.55)*	*(1.83)*	*(2.40)*	*(3.07)*	*(3.62)*
Gender: male[1]	-0.246	-0.349^{+}	-0.691**	-0.396	-0.220	-0.396*	-0.478^{+}	-0.511*	-0.764**
	(-1.20)	*(-1.87)*	*(-2.76)*	*(-1.54)*	*(-1.07)*	*(-1.97)*	*(-1.86)*	*(-2.49)*	*(-2.77)*
Education[2]									
Doctorate	0.702***	0.295^{+}	-0.145	0.619**	0.139	0.0422	0.614*	0.241	-0.237
	(3.57)	*(1.67)*	*(-0.48)*	*(2.70)*	*(0.72)*	*(0.12)*	*(2.55)*	*(1.03)*	*(-0.82)*
High school	0.139	0.0702	0.0376	0.0455	-0.00544	-0.147	0.203	0.180	-0.0297
	(0.58)	*(0.50)*	*(0.14)*	*(0.21)*	*(-0.04)*	*(-0.50)*	*(0.80)*	*(1.03)*	*(-0.08)*
$<$ High school	0.00201	-0.232	-0.0986	0.0447	-0.194	-0.264	0.0653	-0.0885	-0.214
	(0.01)	*(-0.91)*	*(-0.31)*	*(0.22)*	*(-0.84)*	*(-1.45)*	*(0.27)*	*(-0.31)*	*(-0.56)*
Other	-0.113	-0.0766	-0.184	-0.365	-0.407	-0.0393	-0.202	-0.115	-0.130
	(-0.30)	*(-0.15)*	*(-0.23)*	*(-0.61)*	*(-0.70)*	*(-0.03)*	*(-0.39)*	*(-0.25)*	*(-0.12)*
CS background	-0.390**	-0.482***	-0.362**						
	(-2.91)	*(-3.94)*	*(-2.88)*						
Privacy behavior index[3]				-0.0308	-0.0858**	-0.102***			
				(-1.01)	*(-3.06)*	*(-4.82)*			
Westin index							0.152	0.00252	-0.110
							(1.43)	*(0.03)*	*(-0.56)*
Constant	1.466***	2.456***	3.507***	1.438***	2.634***	4.079***	1.061***	2.073***	3.486***
	(6.06)	*(11.22)*	*(9.51)*	*(3.89)*	*(9.28)*	*(9.65)*	*(3.54)*	*(7.24)*	*(6.93)*
Observations		1,031			1,031			1,021	
Pseudo R^2	0.0220	0.0237	0.0259	0.0189	0.0236	0.0303	0.0179	0.0173	0.0239

t statistics in parentheses; $^{+}$ p $<$ 0.10, * p $<$ 0.05, ** p $<$ 0.01, *** p $<$ 0.001; [1](ref. female); [2](ref. university degree); [3]range 0-13

Table 5. Simultaneous quantile regression on the median of weeks each user was trackable using feature set $\mathcal{T}_{\mathrm{u}}^{T}$

Compared to a standard linear regression, a quantile regression provides a "*more complete picture [...] about the relationship between the outcome y [dependent variable] and the regressors x [independent variables] at different points in the conditional distribution of y*" [18]. Thus, it is possible to figure out if there are different effects of the independent variables at different stages of the dependent variable, more precisely: the first quartile (25%), the second quartile (50%, median), and the third quartile (75%). The simultaneous quantile regression estimates three regressions simultaneously (one per quartile) and calculates the standard errors via bootstrapping. Standard errors are used to estimate the t-test for each coefficient and decide whether relations are significant. "*The bootstrap generates multiple samples by resampling from the current sample*" [18] which is necessary to account for multiple testing due to the three simultaneous regressions to ensure correct standard errors and thus correct t-tests.

To analyze the effects of the regressors on the three stages, regression coefficients are compared. Coefficients can be interpreted as changes in the dependent variable, when the independent variable increases by one. As a measure of fit, Pseudo R^2 values are provided. The range of Pseudo R^2 is from 0 to 1, with high Pseudo-R^2 indicating a good fit of the regression model.

Table 5 presents results of three regression models. All models include age, gender and education level, in addition *Model 1* contains the variable *computer science (CS) background*, *Model 2* the *privacy behavior index* and *Model 3* the *Westin index*. These three variables were tested in different models because of the correlation between them. For example, users with a CS background have, on average, a 1.7 points higher privacy behavior index than those without CS background.

Model 1. According to $H1$, there is a relation between age and the number of weeks that users were trackable. The positive coefficients of age indicate that older users are longer trackable than younger ones. For example, the coefficient of 0.0147 in the second quartile means that with each year a person is older, their median trackability is 0.0147 weeks higher, so older users are longer trackable than younger ones.

The t-tests show that the age effects are only significant for the second and third quartile. The age coefficient of 0.0245 means that with each year a person is older, the third quartile of trackability is 0.0245 weeks higher. The slightly higher coefficient indicates that in

the group of long trackability duration, older users are longer trackable.

For gender ($H2$), a negative significant effect can be found for the third quartile. Thus, in the group of the longest trackable users, men are shorter trackable than women with a difference of 0.691 weeks. Significant effects for education ($H3$) can only be found between users with a PhD and a university degree for the first quartile. The first quartile of trackability is 0.7 weeks lower for users with a PhD than for those with a university degree.

According to the three significant negative coefficients in *Model 1*, users with a CS background ($H4$) are shorter trackable than those without: The duration differs from 0.39 to 0.482 and 0.362 weeks, concerning the first, second and third quartile.

Model 2. Here, the same variables are contained as in *Model 1* except for CS background, and the privacy behaviour index (H5) is considered additionally. Age, gender and education show the same effects as in *Model 1*. The coefficients of privacy behaviour index indicate that users differ in the second and third quartiles where those who exhibit more privacy-protecting behaviour are shorter trackable. The difference is 0.086 in the second and 0.102 in the third quartile.

Model 3. Besides age, gender and education, the Westin index ($H6$) is contained. There is no significant relation between the Westin index and trackability.

Summary. There are significant relations between trackability and age, gender, education level, CS background and privacy behaviour index. However, these effects are rather low: The independent variables influence the trackability duration only by some days. Also, the Pseudo R^2 values as the measures of fit are quite low. To sum up, we could find some significant effects, but they only have low impact on the trackability duration.

6 Users' Perspective

Between December 27, 2017 and January 24, 2018, while the study was running for almost two years, we conducted a second user survey (Sec. 3.1.5) to answer the research question RQ4 about users' perception of browser fingerprinting, applied countermeasures and impact of our study.

We recruited participants by sending an email to 760 study participants who were subscribed at that point in time, and 243 (32%) responded.

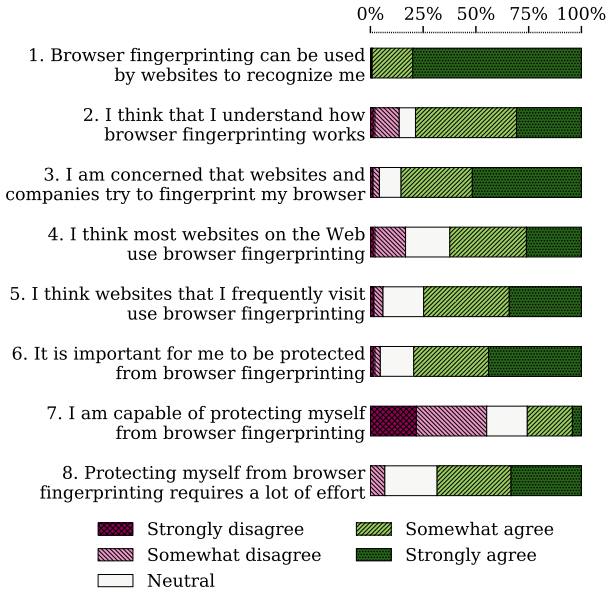

Fig. 4. Agreement w/ statements on browser fingerprinting (N=234)

We performed a qualitative content analysis [19] of the free-text answers about the study impact. For each of the three questions, two researchers read the first 50 answers independently and identified categories for those answers. They then discussed their categories and compiled a codebook together before starting their initial coding for the first 50 answers of each question. Afterwards, they calculated *Cohen's Kappa* κ [20] and discussed their initial coding results and disagreements. Then, they agreed on clearer category descriptions and coded all answers of each question. The achieved intercoder agreement was *excellent* ($\kappa > 0.75$) for 16 out of 22 categories, and *good* ($\kappa > 0.4$) for the remaining 6 categories [21]. Finally, remaining disagreements were discussed, so that full agreement could be reached.

6.1 Understanding, Concern, Protection

234 participants indicated their agreement or disagreement with 8 items (Fig. 4). 99.1% agreed that browser fingerprinting can be used by websites to recognize users. 79.5% indicated that they think they understand how browser fingerprinting works.

85.5% were concerned that websites and companies try to fingerprint their browser. 62.4% thought that most websites use browser fingerprinting, and 74.8% thought that the websites they frequently visit use it.

The majority (78.5%) indicated that being protected from browser fingerprinting is important to them. However, 68.3% agreed that protecting themselves from browser fingerprinting requires a lot of effort, and 55.1% stated they were not capable of protecting themselves.

6.2 New Insights and Experiences

193 participants answered the question whether they experienced or learned something new by participating in our study on browser fingerprinting (see Table 6). 169 of them (87.6%) answered in the affirmative.

"Yes" answers. 75 users (44.4% of 169) stated *technical insights*. For example, P127, learned "*how easy it is to collect user data*", and P151 about "*IP address leakage over WebRTC*", which was one of the features we collected. 64 users (37.9%) experienced their *individual trackability*. Thus, P7 explained: "*It seems I'm easily recognizable when I browse the Internet*", and P29: "*I'm much more recognisable than I thought*".

33 users (19.5%) expressed *awareness*, e.g., P111 learned "*that browser fingerprints exist*" or P91 who learned: "*what fingerprinting actually means and whether it is good or negative*". Finally, 24 (14.2%) of the answers were about *countermeasures*. P38 said: "*[...]Apple's unification [of browser features] results in a worse recognition by websites*", and P67 realized: "*Without blocking JavaScript it is hard to stop fingerprinting*".

"No" answers. 24 participants stated that they did not learn or experienced anything new. 13 of them *did not receive enough information*. For instance, P12 '*[...] knew browser fingerprinting already*", and P98 would have liked "*[...] hints on how to protect myself*". 10 users indicated that they *did not put enough attention or effort* into the study, such as P3 who "*only looked at the [fingerprinting] results*". 10 answers concerned *lack of background knowledge* to understand information that was provided during the study, e.g., P84: "*my knowledge of computers [...] is too low*".

6.3 Participants' Feelings

Did participation in our study changed participants' thoughts or feelings regarding the Web? 166 participants answered this question (see Table 7).

"Yes" answers. 78 participants answered in the affirmative. Most stated that they *realized the consequences* of browser fingerprinting, and thus became disillusioned or disappointed about privacy. Thus, P133 said: "*I did not think that things with user tracking and all this stuff can be so bad. The era of digital human husbandry is here*". Participants also expressed their *insecurity or distrust*, such as P219 who feels "*more persecuted*" or P100 who thinks "*there is real concern that this could be used en masse in the future to track people*".

Some answers referred to *protection*, i.e., wanting countermeasures, or how countermeasures do not exist or perform badly. For instance, P144 expressed: "*Privacy is a real concern nowadays, and I don't think most people can take appropriate measures to protect themselves*". The users also said that they became more *vigilant*, such as P7 who tries to stay offline more often.

"No" answers. 68 of 88 users whose thoughts and feelings did not change were already aware of data collection, browser fingerprinting, or tracking, and thus experienced *no surprises*. Thus, P71 was aware of the "*sad state of internet tracking*" and P80 says that "*companies have always done everything to the disadvantage of their customers*". Some participants expressed their *helplessness* because they cannot escape data collection, whereas other users stated that they have *no problem with tracking*, because "*Advertising finances many free services and I don't want to pay for everything*" (P102).

6.4 Participants' Behavior

We also asked participants whether they changed their behavior on the Web (see Table 8).

"Yes" answers. Out of 155 users, 53 said "yes". Most participants either *applied specific countermeasures* or changed their browsing behavior (*conscious browsing*). For example, P33 uses multiple profiles, P45 "*started to use Firefox for social media/websites that do a lot of tracking, alongside Google Chrome*"; P117 avoids "*unnecessary browsing*", and P4 started to "*block more [content] to browse more safely*".

Some users started *looking for protection*, i.e., either doing research or thinking about it, but not applying countermeasures yet. Others became *more cautious*, such as P143: "*I'm much more cautious with my data and closed my Facebook account after using it for 10 years*". Some users were *unspecific* about their behavior, just stating that they protect themselves.

"No" answers. Most respondents (102 out of 155) did not change their behavior. The main reason was the *lack of protection*: "*I believe my efforts are in vain and it is really up to the Browser Vendors and Website Hosters to do something.*" (P162). Other users think that they are *already protected*, e.g., they use only "trustworthy"

		Category	n	%	κ	Description
Yes	(N=169)	Technical insights	75	44.4	0.68	Learned/experienced how fingerprinting works (details, statistics)
		Individual trackability	64	37.9	0.81	Affected by fingerprinting (e.g., uniqueness, recognizability, stability)
		Awareness	33	19.5	0.73	Gained/raised awareness on fingerprinting
		Countermeasures	24	14.2	0.84	Existence, effectiveness, or importance of countermeasures
No	(N=24)	Not enough information	13	54.2	1.00	Wished or expected more information (e.g., on fingerprinting details)
		Not enough attention/effort	10	41.7	1.00	Did not pay enough attention to study details
		Not enough knowledge	10	41.7	1.00	Lack of background knowledge to understand provided information

Some answers were assigned to multiple categories

Table 6. Answer categories: *Did you experience or learn something new by participating in our study?* ($N = 193$)

		Category	n	%	κ	Description
Yes	(N=78)	Realized consequences	47	60.3	0.76	Disappointment about privacy (e.g., realized magnitude of tracking)
		Insecurity & distrust	19	24.4	0.77	Feeling insecure, worried, distrustful, persecuted, or uncertain
		Protection	14	17.9	0.83	Countermeasures do not exist or perform badly
		Vigilance	12	15.4	0.79	Became (more) cautious while browsing the Web
No	(N=88)	No surprises	68	77.3	0.88	Already aware of data collection, browser fingerprinting, or tracking
		Helplessness	7	8.0	0.82	Helplessness towards data collection and tracking
		No problem with tracking	7	8.0	0.71	Tracking has beneficial aspects, or is not important

Some answers were assigned to multiple categories

Table 7. Answer categories: *Did your study participation change any of your thoughts or feelings regarding the Web?* ($N = 166$)

websites (P16), or do not click on "dangerous links" (P141). Finally, some users did not change their behavior due to the *cost of protection*. P219 stated that "*the most effective ways to protect myself from fingerprinting are also quite invasive.*" and P94 is "*not sure how I could make my browser less unique while still continuing to use my browser for the things I want to do*".

6.5 Applied Countermeasures

We asked participants whether they tried to protect themselves from browser fingerprinting, and if yes, which countermeasures they did apply ($N = 118$). Below, we discuss categories of countermeasures that were applied by more than 5% of the users. We note that many users reported measures that do not protect from fingerprinting, although they are generally good security and privacy practices (see discussion in Sec. 7).

Browser extensions. 38 respondents reported reasonable browser extensions such as *NoScript* or *uBlock* to reduce their fingerprintable features by blocking scripts on specific domains or in general. 10 respondents reported partly effective browser extensions such as *Privacy Badger* which detects and blocks canvas fingerprinting from third-party domains. 26 respondents reported extensions that provide no fingerprinting protection, such as *HTTPS Everywhere* or *CookieAutoDelete*.

Browser settings. Two participants hardened their browser by customizing its configuration (e.g., `privacy.resistFingerprinting` in `about:config`), and 19 participants reported browser settings that do not prevent fingerprinting, such as *Do-not-track*, or disabling third-party cookies.

Disabling JavaScript. Nine respondents disabled JavaScript either completely or partly for some websites and thus limited fingerprinting to HTTP features.

Private browsing modes. Eight participants used the *incognito* or *private browsing mode*, which does not protect from fingerprinting.

Spoofing. Six respondents spoofed their user-agent to hide their genuine browser and operating system. Unfortunately, spoofing can be detected and introduces inconsistencies that might make users distinguishable [14].

Tor browser. 22 respondents used the Tor browser, which is reportedly the currently best countermeasure [22] due to its unification approach and hardening.

Multiple browsers and devices. 22 respondents used different browsers/devices for specific tasks. Although this may separate identities exposed to websites, it does not prevent browser fingerprinting *per se*.

Virtual private network. 10 respondents used VPNs for browser fingerprinting protection. However, this only hides the client's IP address, but is ineffective against browser fingerprinting.

		Category	n	%	κ	Description
Yes	(N=53)	Device-bound protection	20	37.7	0.96	Specific countermeasures (e.g., multiple devices/browsers/profiles)
		Conscious browsing	14	26.4	0.74	Changed browsing behavior (e.g., not visiting certain websites)
		Looking for protection	9	17.0	0.71	Looking for or thinking about countermeasures
		More cautious	7	13.2	0.85	Being more cautious wo/ mentioning specific behavior
		Unspecific protection	7	13.2	0.57	Applying countermeasures or changing behavior but wo/ any details
No	(N=102)	Lack of protection	37	36.3	0.87	Not knowing how to protect oneself / no (proper) defense
		Already protected	27	26.5	0.92	Already cautious enough, or already applying countermeasures
		Cost of protection	16	15.7	0.89	Not changing behaviour due to cost-benefit imbalance

Some answers were assigned to multiple categories

Table 8. Answer categories: *Did your study participation change your behavior on the Web?* ($N = 155$)

7 Discussion

Formal Concepts. Based on the definitions we proposed (Sec. 2.2), a fingerprint is considered trackable if it is unique-by-entity and stable. An entity is considered trackable if it has at least one trackable fingerprint. Although the focus of this paper is on fingerprinting, which is a stateless tracking technique [10], these definitions can also apply on stateful techniques. Tracking cookies, for example, must be unique-by-entity and stable to track individual entities over time.

In practice, adversaries may combine stateless and stateful tracking techniques when they need a certain level of ground truth. Furthermore, a unique stateful identifier can make up for a fingerprint that is (temporarily) not unique. If the goal of an adversary, however, is to distinguish different types of users, such as users with a specific operating system, specific browser language, or users that have visited a specific website at least once, a fingerprint or an identifier does not have to be unique-by-entity. Even if users might not be trackable individually due to being in an anonymity set of size >1 [11], they remain distinguishable from others by being in this particular anonymity set.

Users' Trackability. Since our study enabled participants to use multiple browser and devices, we assessed their trackability using the average and maximum stabilities of their fingerprints (Sec. 4).

Using feature sets from prior works [1, 4] on our test set and considering splits for measurements of desktop and mobile devices as well as all measurements, 67.6%–93.1% of the users were trackable. The mean stability of trackable fingerprints per user averaged to 3.1–3.4 weeks, and the maximum stability to 6.8–8.6 weeks.

Utilizing feature set optimization and crafting device-dependent and device-independent feature sets aiming for stability, we increased the mean stability per user averaged to 10.7–11.9 weeks, and to 20.2–27.2 weeks for the most stable fingerprint per user with 64.6%–94.5% of the users being trackable.

Likewise, we confirmed the applicability of data-driven feature selection on the dataset of Fp-Stalker.

Although the level of ground truth on our dataset is on user level, participants' alternating use of, for example, desktop and mobile browsers should have no qualitative impact on our results since such fingerprints are distinguishable even after stemming features, such as the user-agent string, and thus unlikely to be merged.

Our feature selection increases users' trackability, which has similar negative impact on privacy as the linking of fingerprints that change over time [1, 3]. We believe that both data-driven feature selection and fingerprint linkability should be considered in privacy assessments, as their combination provides a more realistic view on users' trackability.

Although we found significant relations between the trackability of users and their demographics and privacy behavior, the effect sizes were quite low. We did not find any correlations between users' trackability and their use of countermeasures. As our user sample is biased towards German tech-savvy, well educated male users, these relations should be re-examined on user samples that better represent the Internet population.

Device-dependent Fingerprinting. As shown in Appendix C, some selected features (e.g., audio sample rate, accept-language) are similar for both desktop and mobile devices. Other features, e.g., based on Flash, were only picked for desktop devices due to lack of support on mobile devices. We argue that device-dependent fingerprinting is a reasonable strategy for trackers. In practice, such device-dependent feature sets require device type detection. This can only to a limited extend be counteracted by spoofing browser features, as spoofing may yield characteristic inconsistencies [14].

Feature Set Optimization. We showed on two datasets the applicability of feature selection (Sec. 4.2). However, our approach does not yield a one-fits-all feature set for every dataset. Its greedy algorithm may yield suboptimal results. Yet, we believe that the results enable new insights to the importance of features towards certain metrics on a given sample, such as the stability of fingerprints. The greediness of our method can be relaxed by not only yielding the first-best feature candidate for propagation, but all candidates of the same quality while maximizing (or minimizing) towards a given criterion and propagating them tree-wise. Exhaustive, non-greedy propagation could help to assess the (un)importance of features, and pre-filtering feature candidates supports the selection.

Feature Coverage. When we designed our study in 2015/16, the order of HTTP headers [4] had not been discussed in related work yet and could thus not be considered in our evaluation in Sec. 4.2.1. We did not add new features during the ongoing study to have a consistent superset of features. However, we argue that this feature is unlikely to change the results significantly due to its low entropy [4] and low feature importance [3].

Data Collection. Most studies on browser fingerprinting are not performed in a real-world scenario without users' knowledge; thus, users behave differently [5]. Nonetheless, we have thoroughly filtered our data to reduce possible effects on our evaluations induced by participants' curiosity on becoming non-distinguishable.

The availability of data is essential for privacy research on browser fingerprints. Our participants could opt-in to contribute their measurements to a dataset for scientific purposes. This data, however, will be restricted to academia only (upon request) due to the abuse potential by parties not committed to users' privacy.

Users' Perspective. Participants in our study understood that browser fingerprinting can be used to track them, and the majority indicated that they understand how it works. Nevertheless, although we asked participants explicitly about protection against browser fingerprinting (Sec. 3.1.5), a noticeable number of reported protection measures were ineffective against fingerprinting. This might indicate misconceptions about fingerprinting. However, there is also a reasonable chance that participants interpreted this question more broadly in terms of protecting their online behavior in general, and thus reported countermeasures against other forms of threats, such as stateful tracking.

The majority of users were concerned about browser fingerprinting, but they overestimated its prevalence on websites. They expressed that protection is important to them, but that it requires a lot of effort and they do not feel capable of protecting themselves.

The most reported impact of our study on users was awareness of how browser fingerprinting works and their individual trackability. Our study had no impact on most users' behavioral changes, as they reported the lack of protection, already being cautious, or the cost-benefit imbalance of protection. On whether our study affected how users now feel or think about the Web, the respondents were divided: Most who affirmed it are now disappointment about the prevalence of tracking, and most who negated it were already aware about it.

8 Conclusion

In this paper, we presented the results of a 3-year online study on browser fingerprinting with 1,304 participants. Our study is the first to establish ground truth on user level for long-term observations with multiple devices per user, and to provide information on the demographic representativeness of the dataset. Based on two user surveys, we studied users' privacy behavior, their perception of browser fingerprinting, the countermeasures they apply, and the impact of our study. We investigated the trackability of users, and proposed feature stemming and feature set optimization to assess fingerprint stability. Compiling device-dependent and device-independent feature sets, our method increased the mean stability of trackable fingerprints per user by factor 3 compared to existing feature sets [1, 4] on our data, and by factor 2 on a public dataset [3]. Ten years after PANOPTICLICK raised awareness, the situation has not changed: Browser fingerprinting remains a threat to users' privacy. Hopefully, the twentieth anniversary will be more enjoyable.

Acknowledgements

We thank Felix Freiling for simplifying the formalization in Section 2.2. We further thank the anonymous reviewers and our shepherd, Paul Syverson, for their thorough and valuable comments. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

[1] P. Eckersley, "How unique is your web browser?," in *Privacy Enhancing Technologies, 10th International Symposium, PETS 2010, Berlin, Germany, July 21-23, 2010. Proceedings*, pp. 1–18, 2010.

[2] H. Tillmann, "Browser fingerprinting - tracking ohne spuren zu hinterlassen," Master's thesis, 2013.

[3] A. Vastel, P. Laperdrix, W. Rudametkin, and R. Rouvoy, "FP-STALKER: Tracking Browser Fingerprint Evolutions," in *2018 IEEE Symposium on Security and Privacy, SP 2018, Proceedings, 21-23 May 2018, San Francisco, California, USA*, pp. 728–741, 2018.

[4] P. Laperdrix, W. Rudametkin, and B. Baudry, "Beauty and the Beast: Diverting Modern Web Browsers to Build Unique Browser Fingerprints," in *IEEE Symposium on Security and Privacy, SP 2016, San Jose, CA, USA, May 22-26, 2016*, pp. 878–894, 2016.

[5] A. Gómez-Boix, P. Laperdrix, and B. Baudry, "Hiding in the Crowd: An Analysis of the Effectiveness of Browser Fingerprinting at Large Scale," in *Proceedings of the 2018 World Wide Web Conference on World Wide Web, WWW 2018, Lyon, France, April 23-27, 2018*, pp. 309–318, 2018.

[6] D. Fifield and S. Egelman, "Fingerprinting Web Users Through Font Metrics," in *Financial Cryptography and Data Security - 19th International Conference, FC 2015, San Juan, Puerto Rico, January 26-30, 2015, Revised Selected Papers*, pp. 107–124, 2015.

[7] J. R. Mayer, "Any person... a pamphleteer: Internet Anonymity in the Age of Web 2.0," 2009. Bachelor's thesis: https://jonathanmayer.org/publications/thesis09.pdf, accessed on August 5, 2019.

[8] Y. Cao, S. Li, and E. Wijmans, "(Cross-)Browser Fingerprinting via OS and Hardware Level Features," in *24th Annual Network and Distributed System Security Symposium, NDSS 2017, San Diego, California, USA, February 26 - March 1, 2017*, 2017.

[9] A. Cooper, H. Tschofenig, B. Aboba, J. Peterson, J. Morris, M. Hansen and R. Smith, "RFC 6973: Privacy Considerations for Internet Protocols," 2013. https://tools.ietf.org/html/rfc6973, accessed on August 7, 2019.

[10] J. R. Mayer and J. C. Mitchell, "Third-Party Web Tracking: Policy and Technology," in *IEEE Symposium on Security and Privacy, SP 2012, 21-23 May 2012, San Francisco, California, USA*, pp. 413–427, 2012.

[11] C. Díaz, S. Seys, J. Claessens, and B. Preneel, "Towards Measuring Anonymity," in *Privacy Enhancing Technologies, Second International Workshop, PET 2002, San Francisco, CA, USA, April 14-15, 2002, Revised Papers*, pp. 54–68, 2002.

[12] P. Laperdrix, B. Baudry, and V. Mishra, "FPRandom: Randomizing Core Browser Objects to Break Advanced Device Fingerprinting Techniques," in *Engineering Secure Software and Systems - 9th International Symposium, ESSoS 2017, Bonn, Germany, July 3-5, 2017, Proceedings*, pp. 97–114, 2017.

[13] A. Vastel, W. Rudametkin, and R. Rouvoy, "FP-TESTER: Automated Testing of Browser Fingerprint Resilience," in *2018 IEEE European Symposium on Security and Privacy Workshops, EuroS&P Workshops 2018, London, United Kingdom, April 23-27, 2018*, pp. 103–107, 2018.

[14] A. Vastel, P. Laperdrix, W. Rudametkin, and R. Rouvoy, "FP-SCANNER: The Privacy Implications of Browser Fingerprint Inconsistencies," in *27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018.*, pp. 135–150, 2018.

[15] P. Kumaraguru and L. F. Cranor, *Privacy Indexes: A Survey of Westin's Studies*. 2005.

[16] J. B. Lovins, "Development of a Stemming Algorithm," *Mech. Translat. & Comp. Linguistics*, vol. 11, no. 1-2, pp. 22–31, 1968.

[17] L. Molina, L. Belanche, and A. Nebot, "Feature Selection Algorithms: A Survey and Experimental Evaluation," in *IEEE International Conference on Data Mining*, pp. 306–313, 2002.

[18] A. C. Cameron and P. K. Trivedi, "Microeconometrics using Stata, revised edition," *StataCorp LP*, 2010.

[19] M. Schreier, *Qualitative Content Analysis in Practice*. Sage Publications, 2012.

[20] J. Cohen, "A Coefficient of Agreement for Nominal Scales," *Educational and psychological measurement*, vol. 20, no. 1, pp. 37–46, 1960.

[21] M. Banerjee, M. Capozzoli, L. McSweeney, and D. Sinha, "Beyond Kappa: A Review of Interrater Agreement Measures," *Canadian journal of statistics*, vol. 27, no. 1, pp. 3–23, 1999.

[22] A. Datta, J. Lu, and M. C. Tschantz, "Evaluating Anti-Fingerprinting Privacy Enhancing Technologies," in *The World Wide Web Conference, WWW 2019, San Francisco, CA, USA, May 13-17, 2019*, pp. 351–362, 2019.

A Demographic Profile

Table 9 provides an overview on the demographics of our study participants within our evaluated dataset.

B Feature Stemming

Table 10 shows examples for *feature stemming* (see Subsection 4.1) where both original feature values and their *stemmed* versions are shown.

Stemming includes deletion of version substrings, sorting of list-like values in alphabetical order, and assignment of IDs for plugins and MIME types using lowercase concatenations of their properties `name`, `file` and `suffix`, `type`, `desc`, respectively, while removing non-alphabetical characters.

	total		male		female		n/a	
	n	%	n	rel. %	n	rel. %	n	rel. %
Data collection								
participants	1,304	100.0	998	76.5	249	19.1	57	4.4
measurements	88,088	100.0	68,968	78.3	16,445	18.7	2,675	3.0
Age								
Under 18	4	0.3	3	0.3	0	0.0	1	1.8
18-29	324	24.8	261	26.2	57	22.9	6	10.5
30-49	524	40.2	412	41.3	108	43.4	4	7.0
50-64	277	21.2	214	21.4	62	24.9	1	1.8
65 and over	109	8.4	90	9.0	19	7.6	0	0.0
n/a	66	5.1	18	1.8	3	1.2	45	78.9
	1,304	100.0	998	100.0	249	100.0	57	100.0
Education								
Less than high school	124	9.5	82	8.2	40	16.1	2	3.5
High school	130	10.0	106	10.6	23	9.2	1	1.8
University degree	731	56.1	580	58.1	144	57.8	7	12.3
Doctorate	114	8.7	86	8.6	26	10.4	2	3.5
Other	34	2.6	24	2.4	10	4.0	0	0.0
n/a	171	13.1	120	12.0	6	2.4	45	78.9
	1,304	100.0	998	100.0	249	100.0	57	100.0
Occupation								
pupil	23	1.8	21	2.1	0	0.0	2	3.5
student	329	25.2	259	26.0	65	26.1	5	8.8
selfemployed	120	9.2	105	10.5	14	5.6	1	1.8
employee	622	47.7	481	48.2	138	55.4	3	5.3
homemaker	12	0.9	4	0.4	8	3.2	0	0.0
pensioner	73	5.6	61	6.1	12	4.8	0	0.0
unemployed	12	0.9	12	1.2	0	0.0	0	0.0
other	30	2.3	23	2.3	6	2.4	1	1.8
n/a	83	6.4	32	3.2	6	2.4	45	78.9
	1,304	100.0	998	100.0	249	100.0	57	100.0
Background								
Computer science	750	57.5	637	63.8	94	37.8	19	33.3
Non-CS	513	39.3	351	35.2	153	61.4	9	15.8
n/a	41	3.1	10	1.0	2	0.8	29	50.9
	1,304	100.0	998	100.0	249	100.0	57	100.0
Browser Fingerprinting								
Knew before	893	68.5	749	75.1	119	47.8	25	43.9
Did not knew before	382	29.3	249	24.9	130	52.2	3	5.3
n/a	29	2.2	0	0.0	0	0.0	29	50.9
	1,304	100.0	998	100.0	249	100.0	57	100.0

Table 9. Demographic profile of participants in our final dataset

C Feature Set Optimization

The feature sets crafted using data-driven feature selection (see Section 4.2) is shown in Table 11: device-dependent and device-independent feature sets, each optimized towards the number of trackable participants, and towards the avg. stability of trackable fingerprints.

D Features

Due to the sheer amount of 305 browser features that were either collected or derived from existing ones, we report them online: https://browser-fingerprint.cs.fau.de/paper/pets-2020/artifacts/.

Feature	Raw value	Stemmed value
User-Agent (HTTP)	Mozilla/5.0 (Linux; Android 7.0; SM-G920F Build/NRD90M) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/59.0.3071.125 Mobile Safari/537.36	Mozilla (Linux; Android; SM-G920F Build) AppleWebKit (KHTML, like Gecko) Chrome Mobile Safari
navigator.languages (JS)	de,en-US,en	de,en,en-US
navigator.plugins (JS)	[{'file': 'npIntelWebAPIUpdater.dll', 'name': 'Intel® Identity Protection Technology', 'desc': 'Intel web components updater - Installs and updates the Intel web components', 'ver': '5.0.9.0'}, ...]	['intelidentityprotectiontechnology/ npintelwebapiupdater.dll', ...]
navigator.mimeTypes (JS)	[{'desc': 'OpenXPS document', 'suffixes': 'oxps', 'type': 'application/oxps'}, {'desc':'XPSdocument','suffixes':'xps' ,'type':'application/vnd.ms-xpsdocument'} ...]	['oxps/applicationoxps/openxpsdocument', 'xps/applicationvndmsxpsdocument/ xpsdocument', ...]

Table 10. Examples for *feature stemming*

Optimization criterion	**Stability of trackable FPs**			**Participants w/ trackable FPs**		
Device type	Desktop	Mobile	Total	Desktop	Mobile	Total
Feature set	$\mathcal{T}_{\mathrm{s}}^{D}$	$\mathcal{T}_{\mathrm{s}}^{M}$	$\mathcal{T}_{\mathrm{s}}^{T}$	$\mathcal{T}_{\mathrm{u}}^{D}$	$\mathcal{T}_{\mathrm{u}}^{M}$	$\mathcal{T}_{\mathrm{u}}^{T}$
flash_avhardware_disabled	○	○	●	○	○	○
flash_language	●	○	○	○	○	○
flash_os_stemmed	○	○	●	○	○	○
flash_screen_resolution_stemmed_ratio	○	○	●	○	○	○
flash_type_touchscreen	●	○	●	○	○	○
flash_version	○	○	○	○	○	●
fonts_count	○	●	○	○	○	●
fonts_flash	●	○	○	●	○	○
fonts_js	●	○	●	●	○	●
http_accept_language	○	○	○	○	●	○
http_accept_language_stemmed	●	●	●	●	○	●
http_donottrack	●	●	●	○	○	○
http_useragent	○	○	○	○	●	●
http_useragent_stemmed	○	●	●	●	○	○
js_adblocker_enabled	●	●	●	○	●	○
js_app_version	○	○	○	○	○	●
js_audio_channeltype_moz_supported	○	○	●	○	○	○
js_audio_samplerate	●	●	●	○	○	○
js_battery_get_supported	○	●	○	○	○	○
js_battery_level	○	●	○	○	○	○
js_canvas_2d_base64	○	○	○	○	●	○
js_canvas_3d_base64	○	○	○	○	○	●
js_cookies_enabled	○	○	●	○	○	○
js_donottrack	○	○	○	○	●	○
js_donottrack_amiunique	○	○	○	○	○	●
js_donottrack_navigator	○	●	●	○	○	○
js_is_mobile	●	○	○	○	○	○
js_language_browser	○	○	●	○	○	○
js_language_system	○	○	●	○	○	○
js_mimetypes	○	○	○	●	○	○
js_oscpu	○	●	○	○	○	○
js_oscpu_stemmed	○	○	●	○	○	○
js_pdf_reader	○	○	○	○	○	●
js_platform	○	○	○	●	○	○
js_screen_devicepixelratio	○	●	●	○	●	●
js_screen_height	○	○	○	○	●	○
js_screen_height_available	○	○	○	○	●	●
js_screen_resolution_avail_wh	○	●	○	●	○	○
js_screen_resolution_avail_whc	○	○	●	○	○	○
js_screen_resolution_stemmed_avail_whc	●	○	○	○	○	○
js_screen_resolution_stemmed_wh	●	○	○	○	○	○
js_screen_resolution_ratio	○	○	●	○	○	○
js_storage_opendb_windows_enabled	○	○	●	○	○	○
js_storage_session_enabled	○	●	○	○	○	○
js_timezone	○	○	○	○	●	○
js_vibrate_supported	●	○	●	○	○	○
js_webgl_version_stemmed	○	●	●	○	○	○
uap_http_browser_family	●	○	○	○	○	○
uap_http_os_family	●	○	○	○	○	○
uap_http_os_major	●	○	●	●	○	○
uap_http_os_minor	○	●	○	●	○	○
uap_js_device_branding	○	○	●	○	○	○

Table 11. Results of greedy data-driven feature selection on our data