MASTER-SEMINAR
HUMAN CENTERED SECURITY AND PRIVACY

MOBILE SECURITY GROUP

RUHR-UNIVERSITY BOCHUM

Cross-Device Tracking

Author:
Maria Kober

Date:
July 16, 2018

hgi MobSec

RUHR UNIVERSITÄT BOCHUM RUB

Contents

1 Introduction

1.1 Motivation

Many people use different devices, like mobile phones, PCs and tablets, to access online resources every day. Often, there is a connection between the user's behaviour on different devices – like searching for product information on a mobile phone, and buying the product on a PC afterwards. This poses a challenge to companies and services that rely on the ability to identify a single user, e. g., to show personalized advertisement or analyze the effectiveness of shown advertisement. Companies use cross-device tracking to overcome this challenge. [1].

In 2017, Zimmeck et al [1] showed that users are possibly tracked across devices on more than 60% of the websites they visit. This shows that cross-device tracking is widely used. However, researchers lack knowledge on how exactly cross-device tracking companies operate.

According to a TRUSTe survey conducted in 2016 [2, 3], more than 90% of the users in the U.S. and U.K. are worried about their online privacy. Only 25-31% of them understand how companies share their personal data, and during the period of one year about 75% changed their online behaviour due to privacy concerns.

This indicates awareness as well as insecurity on the topic of user tracking. This paper gives an overview of cross-device tracking techniques, use cases as well as arising privacy issues and methods how cross-device tracking can be prevented.

1.2 Outline

Section 2 introduces the two main types of cross-device tracking – deterministic and probabilistic cross-device tracking. For each of the two types, techniques are described how tracking services collect data, allowing them to link different devices to a single user. The last part of Section 2 describes data synchronization between tracking companies.
Section 3 presents use cases for cross-device tracking; the most prominent is (targeted) advertising.
In Section 4 privacy issues related to cross-device tracking are discussed.
Section 5 gives an overview on how to prevent cross-device tracking as a user.
Section 6 concludes this paper with a summary.

2 Overview of Cross-Device Tracking and Techniques

This section provides an overview of the two types of cross-device tracking – probabilistic and deterministic cross-device tracking – and how cross-device tracking can be realized. The last part of this section covers data synchronization between tracking companies.

The focus is on techniques for collecting data. Mechanisms and algorithms on how to match the collected data (e. g., machine learning [1]) and how to present and store this data (e. g., in a *Device Graph* or *Consumer Connection Graph*) are beyond the scope of this paper.

2.1 Deterministic Cross-Device Tracking

Deterministic cross-device tracking is usually based on a first-party relationship between the tracking company and the user. A common example of this relationship is a user who has an account for the company's service [1].

The tracking service of the tracking company uses common persistent identifiers like login credentials, email address, name, HTTP cookies and cookie-like technologies to identify a user. When the user logs in to his or her account from multiple devices, the tracking service can connect those devices to the user's identity [1, 4].

2.2 Probabilistic Cross-Device Tracking

Probabilistic cross-device tracking is usually based on a third-party relationship between the user and the tracking company. It is estimated that the accuracy of probabilistic cross-device tracking is as high as 97% [1, 4].

The first step for successful tracking is to uniquely identify a device, e. g., through device fingerprints or cookies. As a second step, data is collected on those devices. The collected data of several devices is compared and searched for shared attributes. These attributes are used to infer a likelihood that the devices are used by the same person [4].

Probabilistic cross-device tracking also makes use of techniques that are used for single-device-tracking, like web beacons, redirect tracking and referer header analysis [5]. This paper does not cover single-device tracking techniques in detail, but focuses on techniques that can be used for both, single-device tracking and cross-device tracking, and techniques that are mainly used for cross-device linkage.

2.2.1 IP Address

Matching IP addresses is one of the most important techniques for cross-device tracking [1].
It can be assumed that devices belonging to the same user repeatedly share the same IP address. For example, a mobile phone shares the same IP address as the PC at the user's workplace (see Figure 1). At home, the mobile phone shares the same IP address as the user's personal tablet. A tracking service can now link the workplace PC with the user's personal tablet through the mobile phone [4].

Although matching IP addresses is important for cross-device tracking, IP addresses do not always allow to reliably distinguish between users. For example, devices of several users share the same IP address when the users are part of the same household or when they use a public wifi access point [1].

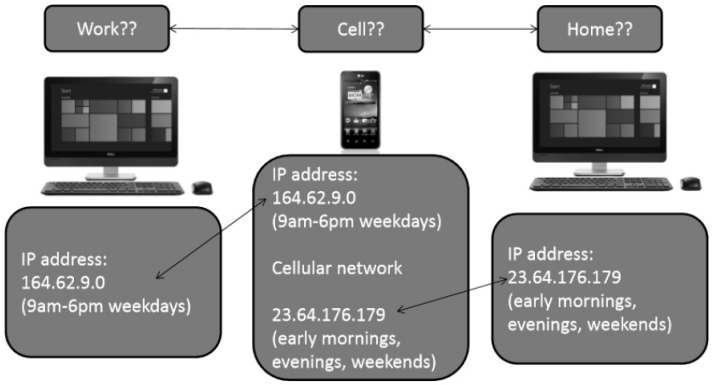

Figure 1: Scheme of IP address matching:
A user uses his mobile phone at work, the mobile phone shares the same IP address as the workplace PC. At home, the mobile phone shares the same IP address as the personal PC of the user. A tracking service can now combine this repeated IP sharing to conclude that all three devices belong to the same user.
Reprinted from [4, Figure 1].

2.2.2 Geolocation Information

The geolocation of several devices can be used to compute a likelihood that those devices belong to the same user. Two or more devices repeatedly sharing the same location can be assumed to belong to the same user or a few users that are related to each other.
Geolocation data is especially helpful when combined with other identifiers, like IP addresses (see Section 2.2.1) [4, 6].

2.2.3 Cookies and Evercookies

Websites can store text files in the web browser of a user. These text files are called HTTP cookies or cookies. They can be used to identify a user's browser, even if the IP address changes. On mobile devices, advertising identifiers, such as Google's Advertising ID, are used in combination with cookie tracking [1].

Zimmeck et al. [1] observed that cookies are used to track users not only on one device and browser, but also across devices. Many websites place not only their own cookies, which can be considered first-party cookies, but also third-party cookies. Malandrino et al. [5] observed that those third-party cookies can make up 80% of the cookies placed by one website.

Evercookies are, like cookies, data files that are stored on a user's device. They are designed to overcome certain limitations of HTTP cookies, particularly they are meant to be more persistent than HTTP cookies. To achieve this, evercookies use not only one but multiple storage vectors like local storage, session storage, and ETags. Some cookies, like Flash Cookies, use

storage that is shared between different browsers. Storage locations are preferred that are less transparent to users and may be more difficult to clear than traditional cookie storage [7].

2.2.4 Device Fingerprints

Device fingerprints are unique, distinguishable and re-identifiable for each device. A fingerprint is based on various device parameters, e. g., API calls to built-in device APIs like the HTML5 Audio API, sensor data and content rendering behaviour. It can be used to bypass tracking- and ad-blocking software [1, 8].

Device fingerprinting includes a variety of techniques. One such technique is canvas fingerprinting (see Figure 2). It can be used to distinguish not only a device but also the used browser. Canvas fingerprinting uses the Canvas API of modern browsers. A tracking service can exploit the rendering behaviour of a device, using always the same content for rendering, e. g., a text in specific font sizes. The rendering behaviour of the browser depends on a variety of system parameters, like the operating system, font libraries, graphic cards and drivers, and the browser itself. This rendering information can be requested by and sent to the tracking service [7].

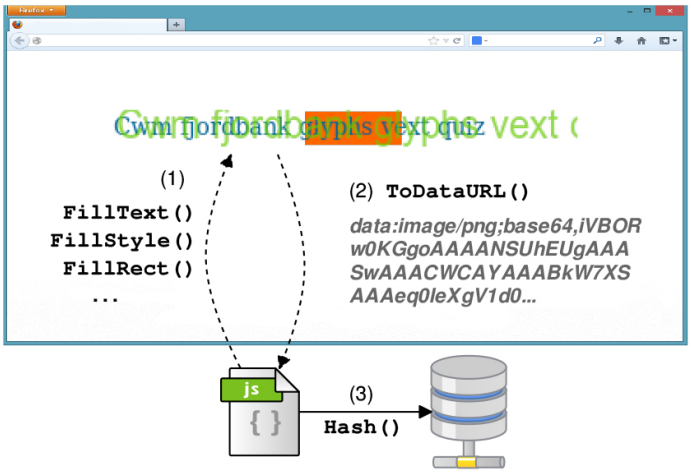

Figure 2: Scheme of Canvas Fingerprinting:
(1) Functions of the Canvas API are invoked, performing the same steps every time the page is loaded. (2) Information that is used for the creation of the canvas fingerprint is requested. (3) The requested data is sent to the tracking service. To reduce data sizes, a hash value can be used.
Reprinted from [7, Figure 1].

2.2.5 Browsing Patterns

The browsing history of a device can be used by tracking services to extract a distinct browsing pattern. Detecting similar browsing patterns on different devices increases the probability that

those devices belong to the same user. In some cases, matching browsing patterns of different devices is a stronger indicator that several devices belong to the same user than IP address matching (see Section 2.2.1) [1, 4].

It is not possible for a website to read out the full browsing history of a user via JavaScript. This means that a tracking service can only track a user's browsing history when it is embedded on the sites the user visits, e.g. as analytic or advertising service. Another way of being embedded on a site is to provide a login capability or a social sharing widget like the *Facebook Like*-button. In this case, the tracking service is treated as a first-party relationship with the user. Embedded third-party services are able to read or place a unique cookie that identifies the device and links it to the first-party site the user is visiting [4, 9, 10].

2.2.6 Ultrasonic Sound Beacons

Several tracking services use sound beacons in the ultrasonic frequency range of $18 - 20$ kHz to track a user. These audio beacons are inaudible to most humans but can be emitted by regular loudspeakers and captured by regular microphones, e. g., the microphone of a user's mobile phone.

Ultrasonic sound beacons rely on the availability of sound emitters. These emitters can be embedded in television commercials, placed in shops or at event locations. Apps embedding appropriate tracking software periodically access the device's microphone to check for ultrasonic frequencies in the device's environment [4, 6, 11].

Using ultrasonic sound beacons enables tracking services to determine the geolocation of a user, independent of GPS signals. It is also possible to track the position of a user within a store by positioning audio emitters at several places within the shop or at the shop entrance. [6, 11]

Different devices repeatedly receiving the same ultrasonic sound beacons indicate that they repeatedly share the same location or are in range of the same TV commercials. This indicates that those devices might belong to the same user.
Furthermore, this tracking technique reveals not only location information about the tracked user but also media behaviour and offline interests [6].

2.3 Data Synchronization

Tracking services, like advertising and analytic services, do cooperate and share data with web and app publishers that have a first-party relationship with the user. This allows more accurate tracking results than the sole usage of probabilistic tracking methods (see Section 2.2) [1].

To identify a user, the first-party identifier itself or the hashed identifier are passed along to third-party tracking services. This allows a tracking service to link the activity of a user on all devices where the same credentials or identifiers are used [4].

Tracking services do collect pseudonymous data, yet it has been observed that personal data is collected, too [7, 12].
According to Mayer [12], there are several ways for third-party services to gather personal data of a user and link it to previously collected (pseudonymous) data:

- The third party is also a first party (see Sections 2.1 and 2.2.5)
- A first party hands off personal data to a third party
- A third party buys personal data
- A third party exploits a security vulnerability in the first party service application
- A third party "deanonymizes" its data by matching it against personal data

Acar et al. [7] observed that data collectors perform server-to-server cookie and data synchronization. This increases the effectiveness of user tracking, especially when combined with other technologies like evercookies (see Sections 2.2.3 and 5.1). However, the authors state that the synchronization of cookies and merging of databases in the background cannot be observed directly. They also state that it is not clear how far cookie synchronization is “common practice” or which companies in particular perform cookie and data synchronization in the background. It has not been observed that different services use the same cookie for the same user. However, using the same cookie value is not necessary for synchronization as different cookies can be mapped to one user in the background [1, 4, 7].

In 2011, Mayer [12] published an empirical study on how personal data of a user is leaked from first-party websites to third-party tracking services. One method to hand over data is the called URL where data like the user’s name, email and country can be stored as parameters at the end of the URL. Third parties embedded into the website will receive this URL, including the personal data. Another way to pass along data from a first party to a third party is the *HTTP Referer header* [8, 12].

3 Use Cases of Cross-Device Tracking

This section provides relevant use cases of cross-device tracking. Most of the presented use cases also apply to single-device tracking.
Several use cases imply privacy issues, which are discussed in Section 4.3. To avoid repetition, some use cases with a focus on privacy implications are presented only in Section 4.3.

Linking data of mobile devices, smart TVs, computers and other devices is used by companies for analytics, including research and testing, account security and advertisement, including ad targeting and purchase attribution (if an advertisement on one device results in a sale on another device) [4].

Zimmeck et al. [1] found that media websites, and in particular websites of newspapers, contain the largest concentration of trackers from cross-device tracking services.

3.1 Advertising and Analytics

Advertising is the main purpose for tracking, including cross-device tracking. Most third-party services that get personal data from first-party services are advertising and analytics services. At least 39% of those tracking services track users across devices and platforms [1, 4, 13].

Advertising services deliver personalized advertising, including targeted and behavioural advertising. Tracking the online activity of a user increases the effectiveness and the revenues of marketing and advertising campaigns. Being able to track a user in real-time also allows the delivery of real-time event-based advertising. This kind of tracking heavily relies on a user's personal data, implying possible privacy issues (see Section 4) [5, 11].

3.2 Search Result Personalization

Search engines, media services, recommendation sites and e-commerce sites like Google, Netflix, Amazon and Yelp provide personalized content. This means they aim to provide content relevant to each single user instead of showing the same content to all users searching for the same query [14].
However, this method is criticized for creating a "Filter Bubble", meaning that the user gets no results on topics, point of views or events that are not in line with the user's previous activities [8].

3.3 Usability Testing

Website developers apply user tracking to test the usability of their websites. Tracking technologies provide tools to observe user actions, the interaction with website elements like buttons, how much time is spent on a particular task, or to highlight issues with the website flow. The collected data can be used to improve the usability of the website in the future [4, 15].

3.4 Account Security

Cross-device tracking can be used to increase a user's account security. This kind of tracking is usually executed by companies that have a first-party relationship with the user.

If someone attempts to log in to an account from a new device, there is a chance that this login is fraudulent. In this case the service the user attempts to log in to may ask for additional authentication. If the user is unable to provide this authentication, the login attempt is seen as fraudulent and he or she is not able to use the device for logging in to the service's account.

Several services and applications provide the user with a list of devices that are currently logged in to his or her account. This list can be used by the user to monitor if there has been an unauthorized access to the account [4].

4 Privacy Issues Related to Cross-Device Tracking

This section covers several privacy issues of cross-device tracking as it is applied nowadays.

Cross-device tracking is a paradigm shift from single-device tracking or browser tracking to the tracking of people. Combining the data collected on several devices and applications, cross-device tracking has the capability to reveal a complete picture of a person, allowing companies to create a comprehensive user profile [1, 6, 16].

Collecting, processing and possibly distributing collected data does raise privacy issues among users that go online. These concerns apply to daily activities like online banking, online shopping and social media interactions [2, 3, 5].

4.1 Data Collection

It is hard to uncover and track how organizations collect personal data from end users, how they store it and if they share it with each other. This is problematic, as not only end users are unaware if they are being tracked but also website publishers and (mobile) app developers might not realize that third parties collect data through their applications [7, 13].

This issue arises as developers use analysis and advertising services that are developed and operated by third parties. Those tracking services obtain the same rights as the application they are integrated into. They may collect user data and share it with third parties, even though the app or website itself is trusted and states that it does not share data with third parties [13].
Two examples for this integration are the *Google* and *Facebook app*. Both are integrated on many websites, giving them the possibility to deterministically collect data on several devices and match the collected data [1].

Ren et al. [16] showed in 2016 that more than 50% of the mobile apps they reviewed leak device identifiers. The device location is leaked by 26% and more than 14% of the apps leaked user identifiers.
In some cases they also observed that data like contact information, location, user identifiers and login credentials are leaked in plaintext. Personal data is also sent to third parties over SSL-secured channels, though the authors state that SSL-traffic contains only a minority of data leaks.

Malandrino et al [5] also observed that tracking companies are able to collect a variety of data, some being personal or confidential data like full name, email address, geolocation information (country, zip-code and city), age and gender, education and employment, health data as well as political and religious beliefs. The authors state that this poses a privacy issue, as, e. g., health information could be combined with personal data and be used to the disadvantage of the user. Another issue noted by the authors is the possibility of identity theft.

Other studies show that there is a relation between the content of a user's email and the advertisement he or she is shown – something a regular user might not expect [1]. Englehardt

et al. [17] observed that third-party tracking services obtain the user's email address and third-party cookies which are stored in the browser when the user views an email. To obtain this data, HTML elements, stylesheets, and embedded images are used in emails.

4.2 Sharing and Selling Data

Razaghpanah et al. [13] observed that the majority of companies they investigated reserve the right to sell or share data with other companies. All of the investigated companies reserve the right to share data with their subsidiaries, meaning that data may be shared even if the company states that it does not share or sell data to third parties.

Companies may sell data including names, postal address, email address, demographic information and behavioural information of users to other companies [8]. According to Bujlow et al. [8] it is common practice for tracking companies to sell collected data to other parties, e. g., insurance companies or online stores.

Malandrino et al. [5] found that companies aggregating collected data exchange this data, including private information. This personal data, like email addresses and credit card numbers, can be linked with pseudo-anonymous data and may be sold by data aggregators. People buying this data might use it for a variety of attacks or illegal actions including identity theft, social engineering attacks and online or offline stalking.

4.3 Data Usage

Several studies showed that collected data is used for targeted advertising (see Section 3.1), implying a variety of privacy issues. Advertisement may be based on sensible data like sexual orientation, financial matters, the health state of a user or email content [8, 18].
However, there are other use cases of tracking that induce privacy issues and user discrimination.

4.3.1 Price Discrimination

Price discrimination is the practice of selling a product to each customer at different prices. The goal of this practice is to sell a good with high profit to customers willing and able to afford a higher price, while also keeping customers who will only buy at lower product prices [19].

Hannak et al. [14] observed that e-commerce sites use collected data to provide different content to each user. This means that users are shown different search results, or that search results are shown in different order. The authors also observed that users were shown different prices, depending on personal parameters like location, if they were logged in to the service and the operating system.

Mikians er al. [19] observed that pricing in e-commerce stores varies depending on the user's location and characteristics. The difference in pricing is based on the personal data that is collected by various online tracking services. The authors found that the pricing difference for the

same product can be as high as three times the price of the cheapest offer. Depending on the geolocation of the user, prices vary from country to country or even between cities.

4.3.2 Financial Creditworthiness

Several online credit institutions use personal data to determine a person's creditworthiness. They use social data, like *Facebook* friends and *Twitter* behaviour, and data achieved through user tracking for their computations. Even more traditional credit institutions, like *American Express*, use online behaviour and social media data to calculate a credit score and determine the creditworthiness of applicants [8, 20].

This method is criticized as not necessarily being indicative whether the applicant will pay back his or her loan [20].

4.3.3 Insurance Coverage and Risk Analysis

Several insurance companies use data collected by marketing companies to gain more information about applicants. Apart from general information about lifestyle, interests, hobbies and habits, information about the types of online articles a user reads are taken into account. This data is linked with data the insurance already has, like medical records, and is used to analyze risks, e. g., the risk for cancer or accidents [8, 21].

4.3.4 Identity Theft and Stalking

Several authors [5, 8, 22] note the issue of identity theft. A user's identity can be stolen by someone buying personal data from a company that is collecting data. Another source of data is information posted by users, e. g., on social media.
There is the danger of online or offline stalking, especially when using publicly available data from social media platforms. This would mean targeted tracking of a single user performed by one individual or a small group of individuals in contrast to being tracked by specialized tracking companies as one of many users [5, 8, 22].

4.4 Transparency

Tracking services provide little to no insight into their business practices. Thus it is not known how they treat cross-device tracking and the data collected through tracking techniques.
Several companies provide the option to opt-out of interest-based advertising. However, it is often complicated and requires following non-standard protocols for a user to opt-out of this kind of marketing. Furthermore, it was shown that opting out does not necessarily prevent user tracking itself (see Section 5.2). [7, 13].

5 How to Prevent Cross-Device Tracking

This section provides information on how to prevent cross-device tracking. It should be noted up front that it is hard to block tracking, including cross-device tracking, in general [7].

5.1 Delete or Block Cookies

To prevent cookie-usage for tracking, users can configure their browsers to delete previously set cookies or block cookies in general. This does not prevent tracking services to use other identifiers, like IP addresses, for tracking. Cookie deletion does not prevent services to reset or respawn cookies (see Section 5.1.1) [4, 7].

Acar et al. [7] state that there is no effective tool to block cookie synchronization in the background. One way to reduce cookie synchronization is to reduce or prevent cookie placement and HTTP traffic as far as possible. However, most users would not want to take such actions as it would reduce or disable the usability of online services.

5.1.1 Cookie Respawning

Cookie respawning describes the re-establishment of HTTP cookies through evercookies, using technology like Flash Storage ("Flash Cookies"), HTML5 local storage and ETags (also see Section 2.2.3). This means that, even though the user deletes browser cookies, they are not permanently deleted (see Figure 3) [7, 23].

Soltani et al. [23] found that an evercookie may re-establish HTTP cookies originated from several sites and not only cookies from the site that placed the evercookie. Using evercookies to respawn deleted cookies enables trackers to link data collected before cookie deletion to data collected after cookie deletion [7, 23].

Disabling evercookie storage may reduce tracking and cookie respawning. However, Acar et al. [7] note that it is impossible to disable certain storage places that are used for storing evercookies, such as localStorage and IndexedDB, without loosing core application functionalities. Additionally the authors found that browser interfaces for deleting cookies are often incomplete or fragmented, resulting in the impossibility to delete all evercookies. Even if the user manages to clear all storage vectors of a browser, cookie respawning is still possible, e. g., through Flash storage which is not isolated but shared between browsers.

5.2 "Do Not Track-Control" and Opting Out of Targeted Advertising

The "Do Not Track"-control is an option that can be set by the user in most modern browsers. It signals that the user does not want third-parties to collect data.

Tracking services do not have to honor the "Do Not Track"-setting, and there are companies that explicitly state that they do not follow it [4]. Several studies [4, 7, 24] showed that the "Do Not Track"-option provides little practical protection against trackers.

Some companies give users the possibility to opt-out of targeted advertising, e. g., companies that are members of the Network Advertising Initiative[1] or the European Interactive Digital Ad-

[1] `https://www.networkadvertising.org/`

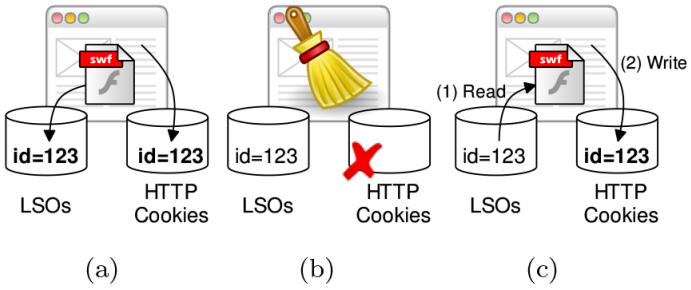

Figure 3: Respawning HTTP cookies by Flash evercookies:
(a) the webpage stores an HTTP and a Flash cookie (LSO), (b) the user removes the HTTP cookie, (c) the webpage respawns the HTTP cookie by copying the value from the Flash cookie.
Picture and description reprinted from [7, Figure 2].

vertising Alliance[2].

Acar et al. [7] observed that neither device fingerprinting nor cookie respawning (see Section 5.1.1) is reduced through opting-out of targeted advertisement. The authors noted that most companies they investigated did not promise to stop the tracking of a user, but only to stop showing targeted advertising. However, the authors observed that opting-out of targeted advertisement does reduce the synchronization of cookies in the background (see Section 2.3).

5.3 Tracker- and Ad-Blocking Software

Tracker- and ad-blocking software blocks domains that are known or believed to perform cross-site tracking. Depending on the specific blocking tool, they also filter or block advertisement and block third-party content and third-party requests.
Some blocking services may whitelist tracking domains that agree to certain standards, such as to honor the user's "Do-Not-Track"-settings (see Section 5.2) [4, 5, 7].

Examples for tracker- and ad-blocking add-ons for browsers:

- AdBlock Plus [3], mainly blocking advertising,
- NoScript [4], blocking JavaScript, Java, Flash and other plugins,
- Ghostery [5], blocking third-party data-tracking technologies,
- RequestPolicy [6], giving users the possibility to control cross-site requests,

[2] https://www.edaa.eu/
[3] https://adblockplus.org/
[4] https://noscript.net/
[5] https://www.ghostery.com/
[6] https://www.requestpolicy.com

- RefControl [7], letting the user set the content of the *HTTP Referer header*.

This paper does not discuss the functionalities of blocking software in detail. A more detailed overview is provided by the evaluation by Malandrino et al. [5] and the overview by Bujlow et al. [8, Table VI].

Blocking third-party connections may not prevent cross-device linkage and -tracking in general. The first-party site the user is visiting can still collect data. This collected data can be matched with data collected across several devices, linked with third-party data or handed over to third-party tracking services [4].

5.4 Tor Browser

By using the Tor browser[8], a user can gain additional protection against tracking and cross-device linkability [4].

Like VPNs (Virtual Private Networks) and anonymous proxy servers, Tor hides IP addresses of devices, successfully preventing IP address tracking (also see Section 2.2.1) [8].

Acar et al [7] note that, in 2014, the Tor browser is the only software they could find that disables canvas fingerprinting, a form of device fingerprinting (see Section 2.2.4).
The Tor browser does so by returning an empty image for all canvas functions that can be used to read image data. The user is then asked whether to trust the site and give it access permission to the canvas object.
The authors also state that the Tor browser appears to be the only effective tool against more traditional fingerprinting techniques. Furthermore, Tor can be used to hide the IP address and geolocation of a user [8].

The Tor Browser Bundle prevents cross-site scripting and third-party cross-site tracking by disabling all third-party cookies and not storing any persistent data [7].

However, Tor does not provide complete security in terms of cross-device tracking. For example, it is possible to uncover the identity of a Tor user through side channels like ultrasonic frequencies [6].

5.5 Privacy-Focused Search Engines

While privacy-focused search engines like DuckDuckGo[9] can not fully prevent user tracking, they do claim that they do not collect any private data. For example, they use methods to prevent search queries from being sent to the sites a user opens through the displayed search results [8].

[7] `http://www.stardrifter.org/`
[8] `https://www.torproject.org/index.html.en`
[9] `https://duckduckgo.com/`

6 Conclusion

Companies and services use cross-device tracking to collect user data. There are two kinds of cross-device tracking, differing in the relationship between the user and the tracking service.

Deterministic cross-device tracking is performed by tracking services that usually have a first-party relationship with the user. Persistent identifiers, like login credentials, email addresses or cookies are used for tracking. When a user logs in to his or her account on several devices, the tracking service can deterministically link those devices and the data collected on those devices together.

Probabilistic cross-device tracking is performed by services that usually have a third-party relationship with the user. These tracking services use a variety of techniques to collect user data on single devices and to compute a likelihood that different devices belong to the same user.
An important technique for probabilistic cross-device tracking is IP address matching, assuming that devices repeatedly sharing the same IP address belong to the same user. Using geolocation information is based on the same principle as IP address matching.
Cookies, evercookies and device fingerprints, unique identifiers based on a variety of device parameters, are also used for cross-device tracking.
Analyzing and comparing the browsing patterns on different devices can help to link devices to a single user, as well as the usage of inaudible audio sounds that are received by device microphones.

It has been observed that tracking companies exchange and synchronize data. This includes pseudonymous data like cookies but also personal data of a user.

The most relevant purpose of cross-device tracking is targeted advertising and analytics. Most third-party tracking services that collect user data are advertising and analytics services.
Another use case is the delivery of personalized content on search engines, e-commerce sites, recommendation sites and the like.
Website developers make use of tracking to test the usability of their websites. Tracking every move of a user can help to find usability issues.
Cross-device tracking is also used to improve account security. When a service receives a login attempt from a “new” device, this login might be fraudulent.

Tracking, including cross-device tracking, implies several privacy issues. Cross-device tracking presents a paradigm shift from device-tracking to user-tracking, allowing companies to create comprehensive user profiles.

The first set of privacy issues concerns the collection of user data. It is difficult to uncover how companies collect personal data, how they store it and if they share it with other companies. This issue also applies to services that are embedded in websites or (mobile) applications. The website publisher or developer might not know which data is collected or shared, and users might be unaware that these applications share data with third parties.

It has been shown that tracking companies collect a variety of personal and confidential data, like health information, level of education, age, religious beliefs and email content. This data can be combined and used in ways that may be disadvantageous for the user.

The second set of privacy issues concerns the practice to share and sell the collected data. It has been observed that several companies exchange, aggregate and sell data, including personal data like names and postal addresses.

The third set of privacy issues concerns the usage of collected data aside from advertisement.
Location information and user characteristics are used to apply price discrimination.
Several credit institutions use collected personal data as well as online and social media behaviour to determine the creditworthiness of an applicant.
Insurance companies receive data about a user's lifestyle, habits, hobbies and other personal data from marketing companies without the user's knowledge. It is used to analyze health risks of applicants.
Personal data can be used for illegal identity theft.

The last set of privacy issues concerns transparency. Tracking companies provide little to no insight into their business practices and thus it is not known how they treat collected user data. Also, several companies provide an opt-out option of interest-based marketing that often is complicated for users to apply to.

Preventing tracking and cross-device tracking is a hard problem.
While deleting and blocking cookies reduces some tracking, deleted cookies can be re-established through evercookies. Deleting cookies does not prevent other types of tracking like IP address matching.
The provided "Do-Not-Track"-control of browsers provides little protection against trackers. Opting out of targeted advertising also does not reduce tracking.
Using tracking- and ad-blocking software prevents some kind of tracking, especially cross-site tracking in browsers, but it does not prevent cross-device tracking and device linkage in general.
Privacy-focused search engines claim that they do not collect any personal data, but they cannot fully prevent user tracking.
Using the Tor browser appears to be one of the most effective tools to prevent several kinds of device and cross-device tracking. However, even the Tor browser cannot prevent tracking completely.

References

[1] Sebastian Zimmeck, Jie S. Li, Hyungtae Kim, Steven M. Bellovin, and Tony Jebara. A Privacy Analysis of Cross-device Tracking. In *Proceedings of 26th USENIX Security Symposium*, 2017.

[2] TRUSTe. U.S. Consumer Privacy Index, 2016. `https://www.trustarc.com/`

resources/privacy-research/ncsa-consumer-privacy-index-us/, as of May 18, 2018.

[3] TRUSTe. GB Consumer Privacy Index, 2016. https://www.trustarc.com/resources/privacy-research/ncsa-consumer-privacy-index-gb/, as of May 18, 2018.

[4] Justin Brookman, Phoebe Rouge, Aaron Alva, and Christina Yeung. Cross-Device Tracking: Measurement and Disclosures. volume 2017, pages 133–148. De Gruyter Open, 2017.

[5] Delfina Malandrino, Andrea Petta, Vittorio Scarano, Luigi Serra, Raffaele Spinelli, and Balachander Krishnamurthy. Privacy Awareness about Information Leakage: Who knows what about me? In *Proceedings of the 12th ACM workshop on Workshop on privacy in the electronic society*, pages 279–284. ACM, 2013.

[6] Daniel Arp, Erwin Quiring, Christian Wressnegger, and Konrad Rieck. Privacy Threats through Ultrasonic Side Channels on Mobile Devices. In *Security and Privacy (EuroS&P), 2017 IEEE European Symposium on*, pages 35–47. IEEE, 2017.

[7] Gunes Acar, Christian Eubank, Steven Englehardt, Marc Juarez, Arvind Narayanan, and Claudia Diaz. The Web Never Forgets: Persistent Tracking Mechanisms in the Wild. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 674–689. ACM, 2014.

[8] Tomasz Bujlow, Valentín Carela-Español, Josep Sole-Pareta, and Pere Barlet-Ros. A Survey on Web Tracking: Mechanisms, Implications, and Defenses. *Proceedings of the IEEE*, 105(8):1476–1510, 2017.

[9] Mozilla. Web technologies for developers > Web APIs > History, October 2017. https://developer.mozilla.org/en-US/docs/Web/API/History, as of May 19, 2018.

[10] Federal Trade Commission. FTC Settlement Puts an End to "History Sniffing" by Online Advertising Network Charged With Deceptively Gathering Data on Consumers, December 2012. https://www.ftc.gov/news-events/press-releases/2012/12/ftc-settlement-puts-end-history-sniffing-online-advertising, as of May 20, 2018.

[11] Vasilios Mavroudis, Shuang Hao, Yanick Fratantonio, Federico Maggi, Christopher Kruegel, and Giovanni Vigna. On the Privacy and Security of the Ultrasound Ecosystem. *Proceedings on Privacy Enhancing Technologies*, 2017(2):95–112, 2017.

[12] Jonathan Mayer. Tracking the trackers: Where everybody knows your username, 2011. https://www.trustarc.com/resources/privacy-research/ncsa-consumer-privacy-index-gb/, as of June 22, 2018.

[13] Abbas Razaghpanah, Rishab Nithyanand, Narseo Vallina-Rodriguez, Srikanth Sundaresan, Mark Allman, Christian Kreibich, and Phillipa Gill. Apps, Trackers, Privacy, and Regulators. February 2018.

[14] Aniko Hannak, Gary Soeller, David Lazer, Alan Mislove, and Christo Wilson. Measuring Price Discrimination and Steering on E-commerce Web Sites. In *Proceedings of the 2014 conference on internet measurement conference*, pages 305–318. ACM, 2014.

[15] Richard Atterer, Monika Wnuk, and Albrecht Schmidt. Knowing the User's Every Move – User Activity Tracking for Website Usability Evaluation and Implicit Interaction. In *Proceedings of the 15th international conference on World Wide Web*, pages 203–212. ACM, 2006.

[16] Jingjing Ren, Ashwin Rao, Martina Lindorfer, Arnaud Legout, and David Choffnes. Recon: Revealing and Controlling PII Leaks in Mobile Network Traffic. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, pages 361–374. ACM, 2016.

[17] Steven Englehardt, Jeffrey Han, and Arvind Narayanan. I never signed up for this! Privacy implications of email tracking. volume 2018, pages 109–126. De Gruyter Open, 2018.

[18] Craig E Wills and Can Tatar. Understanding What They Do with What They Know. In *Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society*, pages 13–18. ACM, 2012.

[19] Jakub Mikians, László Gyarmati, Vijay Erramilli, and Nikolaos Laoutaris. Crowd-assisted Search for Price Discrimination in E-Commerce: First results. In *Proceedings of the ninth ACM conference on Emerging networking experiments and technologies*, pages 1–6. acm, 2013.

[20] Katie Lobosco. Facebook friends could change your credit score, 2013. `http://money.cnn.com/2013/08/26/technology/social/facebook-credit-score/index.html?hpt=hp_t2`, as of June 30, 2018.

[21] Very personal finance, 2012. `https://www.economist.com/finance-and-economics/2012/06/02/very-personal-finance`, as of June 30, 2018.

[22] Ralph Gross and Alessandro Acquisti. Information Revelation and Privacy in Online Social Networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, pages 71–80. ACM, 2005.

[23] Ashkan Soltani, Shannon Canty, Quentin Mayo, Lauren Thomas, and Chris Jay Hoofnagle. Flash Cookies and Privacy. In *AAAI spring symposium: intelligent information privacy management*, volume 2010, pages 158–163, 2010.

[24] Rebecca Balebako, Pedro G. Leon, Richard Shay, Blase Ur, Yang Wang, and Lorrie Faith Cranor. Measuring the Effectiveness of Privacy Tools for Limiting Behavioral Advertising. Web, 2012.