The Creepies and The Crawlies:

Cross device monitoring in web and mobile apps

Max Greenwald
2017
Advised by Professor Arvind Narayanan

Submitted to Princeton University
Department of Computer Science

This thesis represents my own work in accordance with university regulations.

Date of submission: May 5, 2017

Dedicated to humble hard-working men who sacrificed so I wouldn't have to,

Dad '85 and Grandpa '57. And to Mom, a bigger bad ass there never was.

Thank you for teaching me how to hustle.

And finally to my awesome advisor, Arvind, and collaborator, Steve, who stand up each day for the privacy of those who don't know how to protect themselves.

Abstract

Cross-device tracking is a creepy practice where an ad tracking company identifies a consumer via one of their devices and uses that information to identify another of their devices. I performed a novel analysis of 7,561 websites (automated), 16 websites (manual) and 25 iOS mobile apps (manual) to determine the extent of cross-device tracking from a logged-in Facebook user on their devices. Of the websites that had login with Facebook (865 automated and 16 manual), 28.3% of websites (250) and 80% of mobile apps (20) sent plaintext or hashed personally identifiable information (PII) to a third party that was not Facebook. Excluding Facebook, 32 third parties took PII on both a website and mobile app which gives those third parties the potential to conduct cross-device tracking.

Based on the potential harms of cross-device tracking I recommend that the US Federal Trade Commission enact policies to limit the negative effects of cross-device tracking while still encouraging innovation in the space that respects the privacy and security of the consumer. Specifically I advocate that the FTC should 1) encourage that company's privacy policies dictate exactly how and whether cross-device tracking will be implemented 2) work with the DAA to require companies to add good faith single opt-out capabilities from behavioral tracking (and full single opt-out capabilities for top 10 ad space players) and 3) begin a robust education campaign to talk to consumers and importantly, developers of mobile and web applications.

Table of Contents

Introduction

Privacy and security in the digital world are opaque concepts to the average consumer. The last two years has only exacerbated the complexity of how they manifest themselves into our digital lives. In the midst of this opacity, thousands of companies are being created in the advertising and tracking industry. In this industry consumer data is acquired, sold and utilized to offer reduced prices or free services to consumers in exchange for their viewing of increasingly targeted advertisements. This has come at the cost of consumer privacy and security and a new method of consumer tracking, cross-device tracking, is only increasing that cost over time.

Cross-device tracking has taken off in the last 18 months with the advent of mobile tracking. With cross-device tracking a third party company identifies a consumer via one of their devices and uses that information to identify another of their devices. Put together, a company can now serve targeted advertisements to multiple devices of a single consumer to increase their likelihood of purchasing a product. Cross-device tracking provides an encompassing look into a consumer's behavior and has added tracking benefits for a range of purposes, including "ad targeting, research, and conversion attribution" [25]. As this practice proliferates "third parties may use leaked personal information to track app users across multiple websites with knowledge of their real identity" [17]. Furthermore, "sensitive user data may be stored on

badly maintained third-party servers, making them low-hanging fruits for attackers" [17].

Previous methods of consumer tracking that are single device based, specifically placing cookies in the consumer's browser, are no longer as effective for advertisers since "users can switch from laptop to smartphone to tablet an average of 21 times in a single hour" [10]. Elements of effective cookie syncing though can lead to cross-device tracking as first parties "may send cookie values to a cross-device tracking company" and "the cross-device company could return a list of devices it believes to be linked to the same user" [25]. Cross-device tracking services have "the ability to collect richer behavioral and contextual information about users [and] this poses a higher privacy risk than single platform trackers" [19].

Current cross-device tracking policy initiatives including a report by the Federal Trade Commission and new standards by the Digital Advertising Alliance (DAA), a self-regulatory body for the digital advertising industry and enforces responsible privacy practices, have come up short. With the advent of cross-device tracking, Michael Whitener, a data privacy lawyer said that "inevitably, the question is raised whether, in a post-cookie world, a new regulatory regime is necessary to protect privacy" [10].

The key question is the extent to which a tracking company may build a complete profile on consumers' online behavior and create the "database of ruin" a term that means "massive data stores containing hundreds, if not

thousands or tens of thousands, of facts about every member of our society" [36]. What happens when the breadcrumbs of information that each individual consumer leaves on different digital platforms lead tracking companies to put together a complete profile on who, what, when and where consumers browse the web? Various profiles that consumers imagine to be previously separate are now being linked together in new ways. The profiles are also augmented with previously anonymized data sets as tracking companies recognize an individual consumer's behavior. Paul Ohm, a former senior policy advisor for the Federal Trade Commission, said that an exploitable database will allow marketing and tracking companies to "ruin [lives] by the exploitation of data assembled for profit" [36].

Cross-device tracking is creepy, and there needs to be a societal conversation around what kinds of tracking are permissible for companies. A survey done by Pew indicated that "three quarters of internet users are not confident that online advertisers will maintain the privacy and security of their web browsing data" [25]. Consumers deserve to know how companies are using their actively or passively provided information. This thesis explores cross device tracking from a technical and policy perspective. This study examines the extent of cross-device tracking on web and mobile device to get a handle on the tracking ecosystem. Finally, based on the evident potential for cross-device tracking discovered, this thesis makes concrete policy recommendations to the

United States Federal Trade Commission to curb the future harms of cross-device tracking.

Objective Findings

I performed an analysis of 7,561 websites (automated), 16 websites (manual) and 25 iOS mobile apps (manual) to determine which third parties collected identifiers from a logged-in Facebook user on the site. Of the websites that had login with Facebook (865 automated and 16 manual), 28.3% of websites (250) and 80% of mobile apps (20) sent plaintext or hashed personally identifiable information (PII) to a third party that was not Facebook. Excluding Facebook, 32 third parties took PII on both a website and mobile app which gives those third parties the potential to conduct cross-device tracking. There were 9 first party companies where PII was collected on boths its website and iOS app which gives those first parties the potential to conduct cross-device tracking.

Part 1: Cross-Device Tracking Study

Background and prior work

Cross-device tracking has exploded in the last two years because of the advent of connected devices (such as internet of things devices) and the onset

of mobile advertising. John Skovron, the SVP of Platform Engineering at Integral Ad Science, said that “mobile advertising took off literally last year (2016)” and that “most of internet ad spending [by companies] will move from the static web to full format video and in app display on mobile” [22-skovron]. Integral Ad Science, which processes many billions of ad impressions a day, polices first party mobile apps to ensure that third party advertisements are actually being watched by consumers. Ad tracking on mobile is definitely here to stay as ad trackers help make sure that marketers don’t get scammed by mobile apps [22-skovron]. Furthermore it is a lucrative space where a windfall of money is being spent towards improving and augmenting the delivery of mobile ads. Very few studies have tried to get an insight on mobile advertising and tracking, let alone cross-device tracking.

It is important to clarify the nuances within cross-device tracking as there are several types including probabilistic, deterministic logged-in and deterministic shared credential cross-device tracking. Companies can also combine several of these methods in a unique way to accomplish the same goal.

In probabilistic cross-device tracking a company will try to determine the probability that two or more devices are used by the same person by seeing if those devices share any attributes such as an IP address or geolocation. If a phone with a certain IP address is used in two different locations and a computer is only used at one of those locations then it is possible that computer

is a person's work computer while the phone is used at work and home. However if many devices are used using the same WiFi at a coffee shop, this does not meet all of those devices belong to the same user. According to a recent FTC Study, "estimates on the accuracy of probabilistic device correlations range as high as 97.3%. That is, even if users never share identifiers such as an email address or username, companies that use probabilistic device tracking may be able to correctly link devices over 97% of the time" [25].

In deterministic logged-in cross-device tracking a company will take a common persistent identifier (such as a username, birthday or email address) and find it used on several different devices to find the identity of the user. For example a company can put a cookie on the web browser of a computer and then acquire a person's email address through a phone log-in and then tie it back to the cookie on the computer [25]. Google and Facebook are good examples of companies in which a consumer logs on to their services on both web and mobile.

In deterministic shared credential cross-device tracking the tracking companies do not directly interact or have a "login relationship" with consumers. Instead these tracking companies pay or get paid by a first party site (such as Fitbit or Pandora) that has such a direct relationship with consumers. During or after login, the first party site will share those consumer credentials with the tracking companies so they can tie them to other user profiles on different

devices - and on different sites [25]. In all likelihood a consumer will use the same email address on many different services and devices. This study mainly looks at this type, deterministic shared credential cross-device tracking. While many companies perform cross-device tracking it may not be that company's main business. But some services, according to an FTC study, such as Tapad and Drawbridge, are explicitly cross-device tracking companies. Tapad describes itself as "a marketing technology firm renowned for its breakthrough, unified, cross-device solutions " while Drawbridge describes its graph product as "the industry's leading cross-device identity solution, reaching more than one billion consumers across more than five billion digital touchpoints" [25].

While this study adds to current literature in a number of unique ways by combining automated web and manual mobile crawling, previous research papers in the space have studied related topics. A not exhaustive list includes:

- In *Cross-Device Tracking: Measurement and Disclosures*, Rouge et al. did a review of 100 web sites to see which had the potential for cross-device tracking. They found at least 16 out of the 100 sites, "shared personally identifiable information — or hashed personally identifiable information — with third parties, which could allow third parties to correlate multiple devices to persistent real world identifiers" [25].
- In *The Privacy of Just Plain Sites*, Starov et al. looked at 100,000 websites with 30,000 or less monthly views to see how many third parties were present on them. 1500 of these sites has Login With Facebook capabilities and they ascertained which permissions the site asked from Facebook [16].
- In Privacy Leakage vs Protection Measures: the growing disconnect, Krishnamurthy et al. manually looked at PII leakage from the 100 biggest non social sites on HTTP [5].
- In *Are You Sure You Want to Contact Us? Quantifying the Leakage of PII via Website Contact Forms*, Starov et al. looked at 100,000 websites with contact forms to see which leaked PII to third parties [4].

- In *Using the Middle to Meddle with Mobile*, Rao et al. did a study of personally identifiable information (PII) leaking on Android and iOS apps. They found that PII leakage depended on the OS of the device. Of the top hundred apps for iOS and Android, “26 apps [of those surveyed] are available on both iOS and Android. Of these 26 apps, 17 apps leaked PII on at least one OS: 12 apps leaked PIIs only on Android, 2 apps leaked PII only on iOS, while only one app had the same data leakage in both OSes” [28].

Study Methods

This study examines the mobile and web traffic of third party sites present on first party apps and websites to find instances where parties collect PII during and after Facebook login. This study was conducted during Spring of 2017 at Princeton University using the Princeton developed OpenWPM for the web component and Mitmproxy for the iOS component. OpenWPM is a web privacy measurement framework which makes it easy to collect data for privacy studies on a scale of thousands to millions of site [43]. OpenWPM is built on top of Firefox, uses Mitmproxy, and has automation provided by Selenium. Other options one could use for similar technical analysis of the web component could include Janrain with Ajax and PhantomJS with BrowserMob. Other studies have used these however this study used OpenWPM as it is an easy to use open source platform developed at Princeton. Furthermore Janrain costs money and OpenWPM combines the benefits of PhantomJS and BrowserMob. The data collection centered around Facebook login was used because it allowed me to use a fake profile which contained a lot of PII that could be taken, is a typical

action by web and mobile users on a website and finally was a natural point where tracking companies might try to collect PII.

To see if third parties were accessing the identifiers for each app or website, I looked for instances of the fake profiles identifiers being shared. I looked for (1) Device Identifiers specific to a device or OS installation (IMEI, ICCID, iOS IFA and IFV) (2) User Identifiers, which identify the user (name, email address) (3) Location (GPS latitude and longitude, zip code) and (4) Credentials (username, password). These identifiers were chosen to mirror the robust methodology of a previous study [27]. To obtain further coverage of PII leakage I hashed each plaintext identifier using unsalted SHA, Base64, MD5, MMH3, Adler and CRC hashes using a script [40] and also looked for those in the data.

iOS Measurement

On the mobile side I picked a variety of iOS apps with the only criteria being that they have login with Facebook. The selection trended towards those apps that had high privacy sensitivity (fitness, dating) and consumable content (news, music, movie). Table 1 lists the apps examined and their reason for inclusion.

iOS App Studied	Reason For Inclusion
8tracks	consumable content
yelp	privacy sensitivity
cups	consumable content
meetme	privacy sensitivity

espn fantasy	consumable content
tinder	privacy sensitivity
The guardian	consumable content
rec*it	privacy sensitivity
Word Streak	consumable content
cbssports	consumable content
fandango	consumable content
scout	privacy sensitivity
IMDB	consumable content
regal	consumable content
soundcloud	consumable content
bumble	privacy sensitivity
latimes	consumable content
myplate	privacy sensitivity
stumbleupon	consumable content
flashgap	privacy sensitivity
strava	privacy sensitivity
quizlet	consumable content
mapmywalk	privacy sensitivity
shyp	privacy sensitivity
hoteltonight	privacy sensitivity

Table 1: iOS Apps used in Manual Mobile Study

From this set of 25 apps I used Mitmproxy version 0.14 to capture the HTTP and HTTPS traffic of the app as I logged into Facebook and browsed the app for between 45 and 90 seconds to simulate a real user. A fake Facebook profile, Chester Chestnut, was used for each app visited. Several of the apps used certificate pinning, a practice which limits the data I was able to collect. The apps that I was only able to capture some of the data for are 8tracks, Yelp,

Tinder, Bumble, Fandango (only for login), SoundCloud, Quizlet, Rec*It (only for login), Flashgap.

I analyzed the data with a python script [41] using the Mitmproxy depreciated libmproxy library. This takes a mitmproxy flow (version 0.14) and finds all identifiers taken by each 3rd party and puts them into a csv file. The csv file is organized in rows with a third party, identifiers taken by that third party (and if hashed or not), number of PII taken, and a list of hash types used. This code shows how, once an identifier is located in a packet, how to classify it (plaintext or hash) and assign it to the database of it's third party while ignoring duplicates.

```
ids = ['chester', 'other ids', 'md5 hash of chester','other hashes']
numIdentifiers = 18
hashDict = ["md5", "sha1", "sha256", "sha224", "sha384", "sha512", "b64", "crc32",
"adler32", "mmh3", "mmh3-64-1", "mmh3-64-2", "mmh3-128"]
numHashes = 13
database = {}
hashDatabase = {}

if ids.index(id) >= numIdentifiers: ##checks if a plaintext or hashed identifier
    plaintextID = ids[(ids.index(id) - numIdentifiers) / numIdentifiers]
    hashType = hashDict[(ids.index(id) - numIdentifiers) % numHashes]
    plaintextID = hashType + " hash of " + plaintextID
else:
    plaintextID = id
    hashType = ""
if host in database:
    if plaintextID not in database[host]:
        database[host].append(plaintextID)
else:
    database[host] = [plaintextID]
if hashType != "":
    #print hashType
    if host in hashDatabase and hashType not in hashDatabase[host]:
        hashDatabase[host].append(hashType)
    else:
        hashDatabase[host] = [hashType]
elif host not in hashDatabase:
    hashDatabase[host] = [""]
```

The output of this script looks like the Figure 1 below.

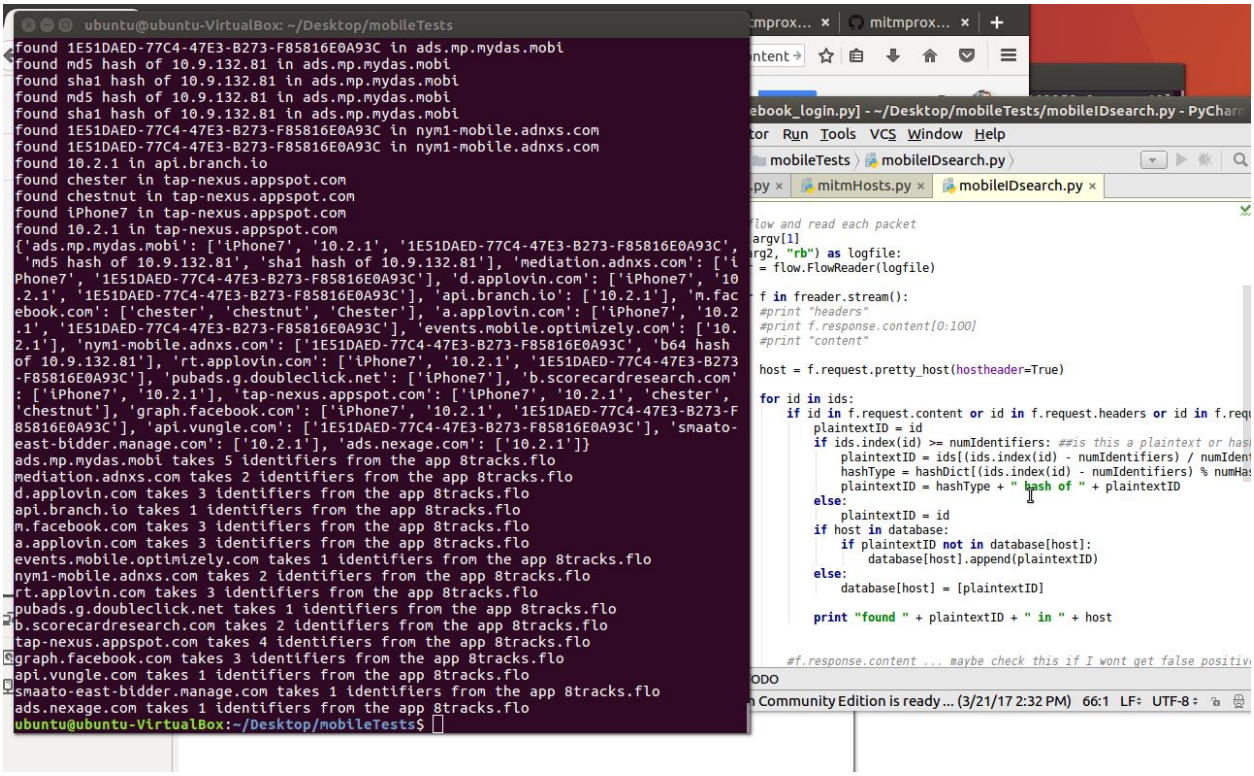


Figure 1: Shows PII taken by 3rd parties operating on the app 8tracks

From there each csv file was converted to Google Sheets where it was cleaned and parsed. Figure 2 shows the raw data output to a csv file for the app MyPlate.

Third Party	Data Taken	# of Data Taken	Types of Hashes Used
graph.facebook.c	['10.2.1', 'iPhone7', '1E51	3	[']
proton.flurry.com	['iPhone7', '10.2.1', '1E51	3	[']
gsp-ssl.ls.apple.c	['iPhone7', '10.2.1']	2	[']
m.facebook.com	['chestnut', 'Chester', 'OU	3	[']
imp.bid.ace.adve	['b64 hash of 10.9.132.81	1	['b64']
data.flurry.com	['iPhone7', '10.2.1', '1E51	5	[']
ssl.google-analyt	['iPhone7']	1	[']
decide.mixpanel.	['iPhone7', '10.2.1', '1E51	3	[']
www.livestrong.c	['chestnut', 'Chester', 'Ch	4	[']
api.mixpanel.com	['b64 hash of 10.9.132.81	1	['b64']
www.googleadse	['1E51DAED-77C4-47E3-	2	[']
sb.scorecardrese	['iPhone7', '10.2.1']	2	[']
api.sessionm.cor	['iPhone7', '10.2.1', '1E51	3	[']
stats.appsflyer.cc	['1E51DAED-77C4-47E3-	1	[']
ad.doubleclick.ne	['1E51DAED-77C4-47E3-	1	[']
a.fiksu.com	['iPhone7', '10.2.1', '1E51	3	[']

Figure 2: Shows PII taken by 3rd parties operating on the app MyPlate

Next I manually examined packets from many of the apps to find more identifiers that I may have missed in the first pass and found 19 worth trying. I realized that my original script was case-sensitive and therefore re-searched using various identifiers with different cases and found more data. Finally I updated the old data with the new results before analyzing it.

Web Measurement

For this component of the study data was obtained using automated and manual analysis. The manual web study was conducted to directly mirror the manual iOS app study to allow for more accurate comparison of the presence of third parties on both devices. The manual web study more accurately simulated

how a user would interact with a site as the automated web study only captured data while clicking 5 random links on each website after a Facebook login. However the automated web study was helpful in expanding the scope of the study to include thousands of websites.

Manual Web Measurement

Just as in the iOS study I used Mitmproxy version 0.14 to capture the HTTP and HTTPS traffic of the app as I logged into Facebook and browsed the website for between 45 and 90 seconds to simulate a real user. A fake Facebook profile, Barley Jenkins, was used for each website. I analyzed the data with a python script [41] using the Mitmproxy depreciated libmproxy library. This takes a mitmproxy flow (version 0.14) and finds all identifiers taken by each 3rd party and puts them into a csv file. The csv file is organized in rows with a third party, identifiers taken by that third party (and if hashed or not), number of PII taken, and a list of hash types used. The code can be examined as explained in the iOS Measurement section.

Automated Web Measurement

The automated interaction of OpenWPM and a Python script located the “sign up” button on a webpage, then proceeded to click the “Log In with Facebook” option, logged in with Facebook and accepted the necessary permissions using Selenium xpath selectors. Finally the crawler checked which third parties received. See Appendix 1 for the Facebook login code. The top 10,000 sites were pulled from the Alexa top 1 million sites

(s3.amazonaws.com/alexa-static/top-1m.csv.zip) though because the server (a c4.2xlarge AWS instance) used did not have enough memory the crawl stopped after 7561 sites. While the site distribution varied in terms of popularity, a previous study showed that the accuracy and reliability of their research did not depend on the site's popularity and that "the distribution of the ranks of the websites where we were successful in identifying and submitting contact forms was uniform" despite the varying crawled sites [4]. Steve Englehardt, a graduate student in the Princeton computer science department, was instrumental in my ability to create and execute this crawler. He both helped advise my code and helped run the crawl.

The crawl worked as follows. The crawl would load each website and visit the homepage. It would then try to login with Facebook. Next it would take a screenshot that could be manually reviewed later. It then re-loaded the homepage and visited 5 links from the homepage while recording all network traffic. Finally it saved this to the sqlite database. Once the data was collected in a sqlite database it was parsed to filter for the sites where a Facebook login page was thought to be detected.

```
SELECT sv.site_url, fb.connect_page_found, fb.connect_successful, fb.fb_api_verified
FROM site_visits as sv LEFT JOIN fb_login as fb ON sv.visit_id = fb.visit_id WHERE
fb.connect_page_found = 1;
```

Figure 3: Shows an SQLite command for information from sites where Facebook login was likely successful

To determine whether a page had actually successfully logged into Facebook I made sure the crawl of a site reached the URL “facebook.com/login.php” and that it entered our fake credentials - this occurred for 865 sites of 7561 (a little over 10%). Other indicators helped to get a rough estimate of the success rate of the Facebook login crawler. First I logged onto the Facebook account of the fake profile and saw that 445 apps had connected with the account (445 of the 865).

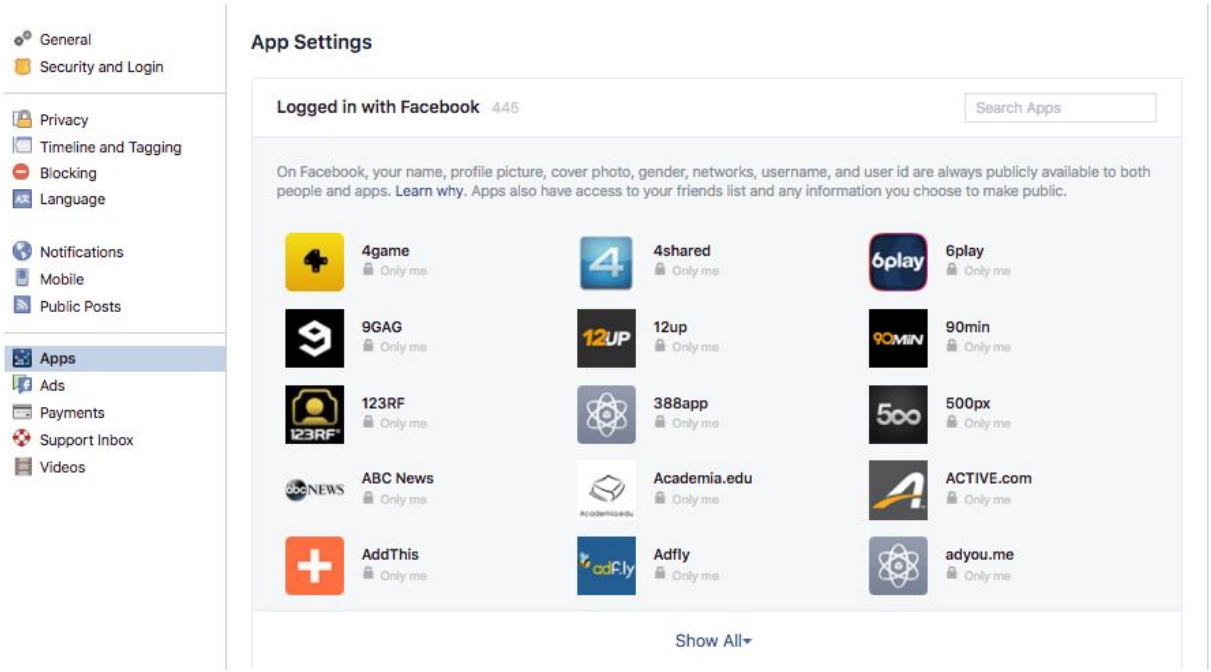


Figure 4: Shows 1st party web apps connected with the fake Facebook account

Next I looked at 150 of the screenshots taken after the Facebook credentials were entered on a site. 100 of the screenshots were of sites that were listed under the Facebook Connected Apps page of the fake account and

50 of the screenshots were of sites that were not listed under the Facebook Connected Apps page. 25 of the 100 screenshots of the Connected Apps had clear evidence that the fake profile was logged into the site. Figure 4 shows clear evidence of the fake profile, Charles, logged into all-free-download.com

Figure 5: Shows successful automatic Facebook login on All-free-download.com

Of the 50 screenshots that were not Facebook Connected, one (Baseball reference) showed a successful login on the screenshot but was not one of the apps officially connected through Facebook. This indicates that for a website that has Facebook login possible and the app does connect with Facebook, the crawler has approximately a 25% success rate. A more detailed explanation is in Appendix C. Extrapolating from the the sample of screenshots, approximately 120 of the 865 sites analyzed completed a full Facebook login process while the

rest suffered from incomplete data. This is to say, the results of the study likely reflect a lower bound for the amount of PII obtained by third parties.

From there I searched in each packet for each first party website (865 sites) for instances of PII or it's hashed values being taken by third parties. This was exported to CSV and analyzed. The full code is available on my Github [44].

Cross-Device Tracking Measurement

To figure out the sites that the potential for cross-device tracking I looked for first and third parties that sent PII on both the mobile and web studies with a simple Python script:

```
for thirdParty in mobilePII:
      if thirdParty in webPII:
            print thirdParty
```

Study Results

The results are broken down between findings from the iOS study, the manual web study, the automated web study and then the combination of all three studies for the cross-device tracking results.

iOS Results

61 third parties were sent PII from the 25 mobile apps studied. 6 third parties were found retrieving PII on more than one mobile app: ad.vrvm.com, api.weather.com, ads.mp.mydas.mobi, api.mixpanel.com,

tap-nexus.appspot.com and api.branch.io. 13 of the third parties sent PII were found to be owned directly by the first party site. Table 2 shows highly sensitive PII sent to third parties on mobile.

App	Third Party	Data Taken
RecIt	api.branch.io	**device fingerprint ID**
Strava	api.branch.io	**device fingerprint ID**
Myplate	data.flurry.com	**name**
Cups	heapanalytics.com	**name**
8Tracks	tap-nexus.appspot.c	**name**
HotelTonight	tap-nexus.appspot.c	**name**
HotelTonight	api.branch.io	**name/device fingerprint ID**
Cups	api.amplitude.com	**name/geolocation**

Table 2: Highly sensitive PII sent to Third Parties on Mobile

Furthermore 26 third parties were sent the geolocation of the mobile phone used in the study. Three of these geolocations were hashed. The most common hash across all sent PII on mobile was SHA1 followed by b64 (an encoding not a hash) and then md5. Table 7 shows some of the identifiers sent hashed to third parties.

App	Third Party	Data Taken	Types of Hashes Used
Myplate	imp.bid.ace.advertising.co	['b64 hash of 10.9.132.i	['b64']
Myplate	api.mixpanel.com	['b64 hash of 10.9.132.i	['b64']
MeetMe	us-east-1.event.prod.bidr.i	['b64 hash of 10.9.132.i	['b64']
LA Times	secure-nym.adnxs.com	['b64 hash of 10.9.132.i	['b64']
LA Times	pixel.adsafeprotected.com	['b64 hash of 10.9.132.i	['b64']
8Tracks	nym1-mobile.adnxs.com	['1E51DAED-77C4-47E	['b64']
8Tracks	ads.mp.mydas.mobi	['iPhone7', '10.2.1', '1E	['md5', 'sha1']
Skout	api.placed.com	['md5 hash of 10.9.132.	['md5']
Skout	ewr-309.ewr-rtb1.rfihub.co	['1E51DAED-77C4-47E	['sha1']
Skout	a.rfihub.com	['iPhone7', '1E51DAED-	['sha1']
Skout	ewr-513.ewr-rtb1.rfihub.co	['1E51DAED-77C4-47E	['sha1']
Skout	haggler-mopub628-us-e-e	['sha1 hash of 10.9.132	['sha1']
Skout	eqv-105.eqv-rtb1.rfihub.co	['1E51DAED-77C4-47E	['sha1']
Skout	ewr-428.ewr-rtb1.rfihub.co	['1E51DAED-77C4-47E	['sha1']
Skout	eqv-160.eqv-rtb1.rfihub.co	['1E51DAED-77C4-47E	['sha1']
MeetMe	haggler-mopub645-us-e-e	['sha1 hash of 10.9.132	['sha1']
MeetMe	haggler-mopub586-us-e-e	['sha1 hash of 10.9.132	['sha1']
MeetMe	haggler-mopub622-us-e-e	['sha1 hash of 10.9.132	['sha1']
MeetMe	haggler-mopub593-us-e-e	['sha1 hash of 10.9.132	['sha1']
LA Times	mpd.mxptint.net	['1E51DAED-77C4-47E	['sha1']
CBS Sports	haggler-doubleclick215-us	['sha1 hash of 10.9.132	['sha1']

Table 3: Hashed identifiers sent to Third Parties on Mobile

The 10 most prevalent third parties, listed in Table 4, were each present on at least 4 of the 25 mobile apps. 36 third parties were present on 2 or more of the apps.

Top 10 Most Prevalent 3rd Parties (4+ Apps)
m.facebook.com
gs-loc.apple.com
haggler-doubleclick215-us-e-ec2.liftoff.io
sb.scorecardresearch.com
ssl.google-analytics.com
ads.mopub.com
ads.mp.mydas.mobi
ads.nexage.com
api.branch.io

app.adjust.com

Table 4: Top 10 Most Prevalent 3rd Parties (4+ Apps)

The raw mobile data is available at bit.ly/CDT-Thesis-Mobile-Data.

Manual Web Results

Only 16 of the 25 apps studied had Login with Facebook capabilities on the web. 33 third parties were sent PII from the 16 sites manually studied. 9 third parties were found retrieving PII on more than one website: Maps.googleapis.com, insight.adsrvr.org, googleads.g.doubleclick.net, sb.scorecardresearch.com, www.facebook.com, www.google-analytics.com, www.google.com, pixel.quantserve.com and geo.moatads.com. 8 of the third parties sent PII were found to be owned directly by the first party site. Table 5 shows highly sensitive PII sent to third parties on mobile.

Website	3rd Party	Data Taken
yelp	d.adroll.com	geolocation/hashed email
shyp	jssdks.mparticle	name/email
mapmyrun	www.google-an	name/gender
fandango	stags.bluekai.co	hashed name/email
latimes	geo.moatads.co	hashed password

Table 5: Highly sensitive PII sent to Third Parties for 16 sites on the Web

Furthermore 5 third parties were sent the geolocation of the computer used in the study. There was also less hashing done in the manual web study

than in mobile with the most used “hash” being b64 (it is an encoding not a hash) Table 6 shows all of the identifiers sent hashed to third parties.

3rd Party	IDs	Number of IDs	Hashs	Website
insight.adsrvr.o	['b64 hash of 14	1	['b64']	regal
geo.moatads.cc	['b64 hash of 14	1	['b64']	regal
px.moatads.cor	['b64 hash of 14	1	['b64']	regal
geo.moatads.cc	['b64 hash of 1(	1	['b64']	latimes
insight.adsrvr.o	['b64 hash of 1(	1	['b64']	latimes
insight.adsrvr.o	['b64 hash of 1(	1	['b64']	espn
geo.moatads.cc	['b64 hash of 1(	1	['b64']	espn
px.moatads.cor	['b64 hash of 1(	1	['b64']	espn
www.regmovie	['1colorado', 'b6	2	['', 'b64']	regal
stags.bluekai.cc	['d2a63577a12(	2	['', 'sha256']	fandango

Table 6: Hashed identifiers sent to Third Parties on 16 sites on the Web

Automated Web Results

The automated web crawl was significantly larger than the manual web crawl (865 sites vs 16). In the automated web crawl there were 173 unique third parties that collected the fake Facebook profile’s first name. 32 first parties sent this name to the 173 third parties (an average of 54 third parties per first party site that sent a name). The first party site that sent the first name the most times was “lesechos.fr” which transmitted the first name of the fake user to 122 first parties while second most, autotrader.com, transmitted the first name to 34 third parties.

147 unique third parties collected the lower and uppercase email address of the fake Facebook user. Third parties login.dotomi.com and pippio.com

collected the email address on 65 different first party sites. Since only 60 first parties leaked a lower or uppercase email at least five first party sites leaked pippio.com both a lower and uppercase version of the profile's email. One third party, sync.graph.bluecava.com, collected the same uniqueID from the Facebook user on 29 different first parties.

Gogoanime.io sent the profile's email address to the most third parties of any of other first party, at 122 third parties. 21 third parties received the user's zip code, while Doubleclick, a Google subsidiary, received both the user's email and zipcode. Table 7 lists the 17 third parties across the automated web study that received the email, first and last name of the fake Facebook user.

Third Parties
zoomus.zendesk.com
securepubads.g.doubleclick.net
beacon.krxd.net
secure.adnxs.com
ib.adnxs.com
pixel.rubiconproject.com
dsum-sec.casalemedia.com
match.adsrvr.org
api-iam.intercom.io
dev.appboy.com
app.satismeter.com
api.segment.io
api.amplitude.com
na.wargaming.net
qatarliving.zendesk.com
www.lyrster.com

eu.wargaming.net

Table 7: Third parties that received the user's email, first and last name

Cross-Device Tracking Results

Looking first at the manual studies (25 iOS apps and 16 websites), there were 9 first party companies that shared PII with third parties on both their mobile app and website.

First Party Cross-Device PII Sharers
strava
yelp
mapmyrun
8tracks
stumbleupon
shyp
espn
regal
fandango

Table 8: First parties that shared PII on both a mobile app and website

Across the manual studies there was 15 third parties that took some identifiers from both a mobile app and website. 7 out of 15 of those third parties took that some identifiers from the same app, while 8 out of 15 took mobile some identifiers from one app and web some identifiers from another app.

3rd Parties that took some IDs from Manual Study	Present On Same App + Web
facebook.com	ALL
beacon.krxd.net	LA Times, Guardin, Fandango
ad.doubleclick.net	guardian, LA Times
api.shyp.com	shyp
m.trb.com	LA Times
registerdisney.go.com	ESPN
stags.bluekai.com	Fandango
pixel.quantserve.com	NONE
px.moatads.com	NONE
d.agkn.com	NONE
dpm.demdex.net	NONE
p.adsymptotic.com	NONE
sb.scorecardresearch.com	NONE
api.amplitude.com	NONE
idsync.rlcdn.com	NONE

Table 9: Third parties that collected some identifiers on both a mobile app and website and which apps they collected it on

In Table 10 there is a list of 4 third parties that collect personally identifiable information from a first party app and first party manually collected website.

Third Parties and the PII they took (from which app)	
pixel.quantserve.com	geolocation (meetme) and name (soundcloud, quizlet, 8tracks)
api.shyp.com	name (shyp) and credentials (shyp)
registerdisney.go.com	name (espn) and credentials (espn)
api.amplitude.com	name (shyp, cups) and geolocation (cups)

Table 10: Third parties that collected PII on both a mobile app and website (and which apps they collected it on)

Across the manual iOS study, the manual web study and the automated web study there were 32 third parties that collected PII on both a mobile app and a website.

Third Parties that collected PII on mobile and web		
aa.agkn.com	geo.moatads.com	sb.scorecardresearch.com
ad.doubleclick.net	googleads.g.doubleclick.net	securepubads.g.doubleclick.net
api.amplitude.com	graph.facebook.com	ssl.google-analytics.com
api.shyp.com	idsync.rlcdn.com	ssp.lkqd.net
as.eu.angsrvr.com	m.trb.com	stags.bluekai.com
b.scorecardresearch.com	p.adsymptotic.com	t.lkqd.net
beacon.krxd.net	pagead2.googlesyndication.com	tags.bluekai.com
bs.serving-sys.com	pixel.adsafeprotected.com	trc.taboola.com
cdn.krxd.net	pixel.quantserve.com	www.googleadservices.com
d.agkn.com	px.moatads.com	www.stumbleupon.com
dpm.demdex.net	registerdisney.go.com	

Table 11: Third parties that collected PII on both a mobile app and website

Table 12 is a comparison between the PII that a third party took from mobile and from web. On the left column is the PII that the third party took from the web while the column on the right is the PII that the third party took from mobile. The short numbers are geolocation, while the names are some of the first and last names of the fake profiles used to collect the data. Though only latitude or longitude is listed for the geolocations, both lat and long were taken by the third party. The IDs on the mobile side are either IDFA's or other identifiers. The numbers with several periods in them are IP addresses. Many third parties took multiple pieces of PII from a single device.

PII Taken on Web	CDT 3rd Party	PII Taken on Mobile
chief_wiggins@hotmail.com	aa.agkn.com	1e51daed-77c4-47e3-b273-f85816e0a93c
40.34	ad.doubleclick.net	1e51daed-77c4-47e3-b273-f85816e0a93c
Wiggins, Barley', 'Jenkins	api.amplitude.com	40.34', 'Chester', 'Chestnut', 'chestnut', 'sha256 hash of chestnut
40.34	as.eu.angsrvr.com	140.180.251.187
Wiggins	beacon.krxd.net	10.2.1', '1E51DAED-77C4-47E3-B273-F85816E0A93C
40.34	bs.serving-sys.com	40.34', '1E51DAED-77C4-47E3-B273-F85816E0A93C
Wiggins	cdn.krxd.net	1E51DAED-77C4-47E3-B273-F85816E0A93C
chief_wiggins@hotmail.com	d.agkn.com	1E51DAED-77C4-47E3-B273-F85816E0A93C
40.34,Barley', 'Jenkins	googleads.g.doubleclick.net	iPhone7
-74.6	idsync.rlcdn.com	1e51daed-77c4-47e3-b273-f85816e0a93c
charles	p.adsymptotic.com	1E51DAED-77C4-47E3-B273-F85816E0A93C
-74.6	pagead2.googlesyndication.(	iPhone7', '10.2.1', '1E51DAED-77C4-47E3-B273-F85816E0A93C
-74.6	pixel.adsafeprotected.com	b64 hash of 10.9.132.81
40.34, barley', 'jenkins	pixel.quantserve.com	iPhone7', '10.2.1', '1E51DAED-77C4-47E3-B273-F85816E0A93C', '40.34
40.34	px.moatads.com	iPhone7
chief_wiggins@hotmail.com	registerdisney.go.com	chestnut', 'chestmchestnut@outlook.com', 'Chester', 'Chestnut
40.34	sb.scorecardresearch.com	1e51daed-77c4-47e3-b273-f85816e0a93c
-74.6	securepubads.g.doubleclick.	1E51DAED-77C4-47E3-B273-F85816E0A93C
40.34	ssl.google-analytics.com	iPhone7
40.34	ssp.lkqd.net	140.180.251.187
40.34, md5 hash of 'herewe	stags.bluekai.com	1E51DAED-77C4-47E3-B273-F85816E0A93C
40.34	t.lkqd.net	140.180.251.187
-74.6	tags.bluekai.com	1e51daed-77c4-47e3-b273-f85816e0a93,c
40.34	trc.taboola.com	1E51DAED-77C4-47E3-B273-F85816E0A93C
40.34	www.googleadservices.com	1E51DAED-77C4-47E3-B273-F85816E0A93C', '10.2.1

Table 12: Breakdown of what PII a third party took on both a mobile app and website

Discussion

Across the mobile and web data credentials, geolocation, device fingerprint IDs, zip code and private IP addresses were shared with hundreds of third parties. Each of these PII alone may not be inherently worrisome for an internet or mobile user today. But in aggregate it might be significantly more concerning.

Based on the results there is an ability for at least 25 companies to engage in cross-device tracking. From the above list note several third parties

are hosts owned by the same company: doubleclick (x2), moatads, bluekai, lkqd, scorecardresearch, agkn - and therefore are not added into the final tally. Adsymptotic, owned by Drawbridge, and adkn, owned by Neustar, are two example of these 25 companies that explicitly discuss intent to use cross-device tracking. Neustar says it “resolves disparate consumer identifiers to ensure that Neustar can recognize them collectively as a single consumer across devices.” While not all of them explicitly describe themselves as conducting cross-device tracking, it is enough to be concerning. Cross-device tracking can also occur on the backend making it hard to know when companies are doing it “since companies can make determinations of device correlation on their own servers, unobservable to end users” [25].

Other types of cross-tracking may be occurring as well. The 36 third parties that took PII on 2+ mobile sites are potentially able to engage in cross-app tracking. Third parties login.dotomi.com and pippio.com collected the user’s email address on 65 different first party sites on web which could allow them to track a user across websites. It also seems concerning that a few first parties, like Gogoanime.io, sent PII to over 100 third party sites.

One interesting comparison to make is the most prevalent third parties on my web study to a recent Federal Trade Commission study.

Table 2. Top 20 third-party domains with most connections from 100 sites tested

Third-Party Domain	Run1	Run2	Run3	Run4	Average
doubleclick.net	88	89	87	86	87.5
facebook.com	69	71	68	68	69
google.com	70	69	70	62	67.75
google-analytics.com	65	67	64	58	63.5
scorecardresearch.com	65	60	61	58	61
googlesyndication.com	62	63	58	58	60.25
adnxs.com	48	47	48	50	48.25
2mdn.net	48	49	44	46	46.75
gstatic.com	49	55	4	34	46
googleapis.com	47	54	38	43	45.5
cloudfront.net	46	48	44	41	44.75
yahoo.com	47	50	44	36	44.25
moatads.com	47	46	42	40	43.75
bluekai.com	44	45	40	39	42
twitter.com	43	41	40	32	39
advertising.com	35	41	42	37	38.75
rubiconproject.com	40	37	38	39	38.5
adsafeprotected.com	38	38	41	35	38
rlcdn.com	34	38	38	37	36.75
imrworldwide.com	37	39	38	32	36.5

[25]

Third Party On Web	Number of Websites Appeared On
pippio.com	67
idsync.rlcdn.com	66
tapestry.tapad.com	66
login.dotomi.com	65
s.thebrighttag.com	65
aa.agkn.com	64
ib.adnxs.com	64
pixel.tapad.com	64
v12group.com	64
stags.bluekai.com	60
api.traversedlp.com	59
p.alcmpn.com	59
p.nexac.com	59
cm.g.doubleclick.net	58
static.traversedlp.com	58
a248.e.akamai.net	56
login-ds.dotomi.com	53
ei.rlcdn.com	49
geo-um.btrll.com	47
pixel.rubiconproject.com	46

Figure 6: Top 20 most prevalent third parties in an FTC cross-device study on web and this study on web

Overlaps include rubiconproject, rcldn, bluekai, adnxs and doubleclick though the FTC rate of prevalence on their 100 sites they looked at is higher than the 865 sites that I looked at - likely because they manually navigated on each site and therefore generated more data than my automatic web crawl which only clicked on 5 links and likely only logged in successfully to 120 sites with Facebook.

My study seems to have found similar top third parties to other studies. In a International Computer Science Institute (ICSI) at Berkeley study called Tracking the Trackers they compared how often top found mobile and web

advertising and tracking services (ATS) appear on the opposite device (with the top Alexa 1000 sites).

ATS Domain	ABP	hpHosts	#Apps	#Sites
crashlytics.com	False	False	434	0
facebook.com	False	True	406	623
doubleclick.net	True	True	190	621
gstatic.com	False	True	172	509
googlesyndication.com	False	True	160	441
flurry.com	True	True	133	0
appsflyer.com	False	True	95	9
google-analytics.com	True	True	95	664
googletagmanager.com	True	True	78	200
googleadservices.com	True	True	72	470

Table 4: Top 10 ATS domains (sorted by app penetration) with their presence on manually curated ATS lists and penetration in the Alexa Top 1000 Websites.

Figure 7: Prevalence of top tracking services are mobile and web

Similar to this study, third parties like crashlytics, flurry and appsflyer tended to appear on mobile while google-analytics and doubleclick.com tended to appear on web. Facebook similarly was present heavily on mobile and web. This could be an indication that Facebook has higher market penetration on both devices for doing tracking and analytics.

The tracking study showed that 68.5% of advertising and tracking services “are cross-platform and operate on at least one website in the Alexa Top 1000.” The most prevalent of the advertising and tracking services - Facebook, DoubleClick, and Google Analytics - are “present on over 60% of all the Alexa Top 1000 websites” [19] whereas mine showed lower rates of only 15-20% presence across websites I looked at (100% for Facebook because attempted log in with Facebook was a prerequisite to qualify for the study). This lowered

percentage is likely due in part to the hampered data of the Facebook crawl, but also that I looked at the top 7561 sites not the top 1000 sites and am specifically looking at sites with a large Facebook presence. It seems likely that sites with a large Facebook presence might not also have a large Google presence (5 of the top 10 most present ATS's in the Tracking the Trackers study are Google third parties).

Error Rates

I acknowledge a potential for false positive or false negative errors with my data. For example there were likely several hashed pieces of PII that I did not find. It is also possible that the values I used to search for geolocation sharing (40.XX and -74.XX where the X's represent varied decimal place searches) might have been short enough that they might get picked up in a data packet when they were actually part of another value. Another study that manually looked at 100 web forms of an automated study found that 58% of the time their automated and manual crawling matched perfectly. In the non perfect matches they found 131 false positives or false negatives which amounts to a 6.24% error rate. This was described as "not perfect but pretty good" [16] and hopefully this study is on par with this error rate.

Part 2: Policy for Cross-Device Tracking

This thesis now shifts gears to examine the current policy surrounding cross-device tracking and what policy options relevant authoritative bodies could enact to mediate in the space.

Policy Framework Background

Policy makers should find better ways to articulate what kind of tracking is and is not acceptable. From web tracking to cross device tracking to cross-app tracking it is unclear what rights consumers have against trackers [30]. There are several players that that could have a hand in cross-device tracking policy. There are two government agencies: the Federal Trade Commission (FTC) and the Federal Communications Commission (FCC). There is the self-policing advertising organization called the Digital Advertising Alliance (DAA). And finally there is the nonprofit Council of Better Business Bureaus (CBBB) which works to accredit businesses for quality standards.

The FTC has moved into the data security and data privacy space under its Section 5 authority to stop deceptive and unfair practices of companies. Over the last 20 years it has litigated over 50 data security and privacy cases and "has sought to bring greater transparency and user control to the issue of online behavioral data collection as part of its work to protect and promote consumer privacy" [25].

The FTC has yet to bring “an enforcement action specifically targeting cross-device tracking, [however] it appears clear that the FTC’s broad authority under Section 5(a) of the FTC Act to prevent “unfair or deceptive acts or practices” would allow it to do so” [10]. It has already had a few cases against tracking companies such as Flash cookies (ScanScout, Inc.) and history-sniffing scripts (Epic Marketplace, Inc.).

As it has became more aware of the behavioral targeting and advertising space it held a behavior targeting workshop in 2007 and published “Self-Regulatory Principles for Online Advertising” in 2009. Behavioral advertising was also a “significant focus of the 2012 Report ‘Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers’” [25].

The other government organization that could act in the cross-device tracking space is the FCC, which is an agency in charge of regulating radio, television, wire, satellite, and cable. This does extend to the internet as well since the FCC was responsible for the 2015 open internet rules. In 2016 the FCC passed new privacy rules for “ISPs to get opt-in consent from consumers before sharing Web browsing data and other private information with advertisers and other third parties” [37]. However this was halted in 2017 by the US Senate so it unlikely these rules will go into effect soon.

The FCC 2016 proposed rulemaking report argues for companies to minimize their data collection. They discuss “data minimization, including

whether [they] should impose reasonable data collection and retention limits. [They] also seek comment on whether [they] should prescribe specific data destruction policies as part of any data retention limits" [21]. The FCC also further questions whether certain types of data should be prohibited from collection all together stating, "Are there particular types of customer data, such as health information, that a provider should be prohibited from collecting?" At the same time the FCC recognizes that this could be difficult for companies by asking, "Could such a requirement be implemented and operationalized without undue burden" [21]?

The next player in the space is the DAA which establishes and enforces responsible privacy practices across industry for relevant digital advertising. The DAA has made clear that "for uses other than certain excepted uses (e.g., intellectual property protection, consumer safety, research, authentication, etc.), and most namely interest-based advertising, participants must provide a consumer opt-out" [34] administered by the DAA through its AdChoices and AppChoices programs.

Their most effective program called AdChoices gives users more transparency and control other their ads and is opt-in for companies. It is recognizable "by its icon [placed on the ad] with more information about the ad or the website's collection practices. Over 60% of ads in a sample of 183 ads from top news websites are covered by AdChoices "[23].

Recently, the DAA offered some guidance on cross-device tracking and updated their own enforcement language to say that companies' choices for ads for a consumer on this device "will apply to data collected…from other browsers or devices"[35] and that "no browsing and usage data may flow into or out of that device/browser for the purposes of internet-based advertising"[35]. This indicates that if a consumer opts out of behavioral targeting on one device, companies cannot target them behaviorally on another device that the consumer is known to have. Compliance to the DAA is often "contractually required by and amongst advertisers, ad agencies, ad networks, and publishers. DAA participants are expected to publicly commit to compliance within its principles" [34]. Cleverly, if a company publicly commits to compliance and then fails to do so, it would count as "a false advertising statement the FTC and state regulators can [then] prosecute as a deceptive practice" [34].

In its guidance on cross-device tracking, the DAA pleaded with companies saying: "let's be sure we keep true to our principles of enhanced transparency and consumer control. The reward is better consumer engagement and confidence through a responsible internet-based advertising ecosystem with meaningful accountability." [35] However just following the DAA guidelines surrounding asking for permission to collect data is not sufficient, "although that is typically the direction that US regulators take (for example the FTC cross-device tracking report recommendations)" [30].

There is also the Council of Better Business Bureau (CBBB). The CBBB is dedicated to fostering honest and responsive relationships between businesses and consumers. It gives negative ratings to businesses who lack these relationships. In response to the DAA's recent guidance, on February 1, 2017 the CBBB began taking complaints and "monitoring the marketplace for first-party and third-party transparency and choice to consumers as they pertain to cross-device data collection" [35].

Lastly one can also look to advertising regulation counterparts in the European Union. These regulators are stricter than in the United States and simply advise against any tracking, which is also a tough call. This thesis will focus on US based policy options for cross-device tracking because "the largest advertising-supported businesses are based in the United States and because policy disputes about advertisement blocking have tended to arise in the United States"[23] before the European courts.

Policy Recommendations

The US Federal Trade Commission should enact policies to limit the negative effects of cross-device tracking while still encouraging innovation in the space that respects the privacy and security of the consumer. Specifically the FTC should 1) encourage that company's privacy policies dictate exactly how and whether cross-device tracking will be implemented 2) work with the DAA to

require companies to add good faith single opt-out capabilities from behavioral tracking (and full single opt-out capabilities for top 10 ad space players) and 3) begin a robust education campaign to talk to consumers and importantly, developers of mobile and web applications.

Privacy Policy Clarification

The FTC should encourage companies to write consumer readable privacy policies that specifically dictate to what extent consumer data will be used and distributed. Particularly in regards to cross-device tracking. Most privacy policies today are dozens of pages long with difficult to read legalise. They are also often intentionally vague to maximize their right to collect data and protect themselves against a future lawsuit. In the FTC web study of 100 popular websites, most of the policies reviewed “reserve[d] broad rights to allow third parties to collect and use pseudonymous browser data such as IP address and unique cookie identifiers” [25]. Companies should be able to update their privacy policies as needed to broaden the data they collect (as long as the consumer is informed) but should be required to start with the minimum data they need and broaden from there.

Companies need to be more specific especially in regards to cross-device tracking since consumers may not understand “the extent of data mining or that anonymous identifiers and hashed personally identifiable information can still be linked to a particular consumer. Further, consumers may not expect that sensitive data could be derived from pieces of data that are not traditionally

sensitive (e.g., websites visited)" [34]. Greater transparency and choices for consumers is essential and can be explained via a privacy policy. In the FTC workshop on cross-device tracking, several panelists argued that "there are few tools that allow consumers to understand which devices are linked to their device graphs" [34] which is something that a privacy policy could clarify.

Opt-Out Capabilities working with the DAA

The FTC should work with the DAA to require companies to add good faith single opt-out capabilities from behavioral tracking (and full single opt-out capabilities for top 10 ad space players). Companies rarely provide the ability for consumers to opt out of behavioral advertising. Where such tools are present, they only allow for opting out of targeted advertising, not cross-device tracking. One panelist at the FTC cross-device tracking workshop suggested that "consumers should be able to opt out of entire device graphs using a single opt-out" [34] which this thesis concurs with. A single opt-out point makes the decision making easier on the consumer as well as minimizes consumer confusion and lowers the knowledge barrier for the three quarters of consumers that are "not confident that online advertisers will maintain the privacy and security of their web browsing data" [25] but do not have the technical specificity to understand how to easily take action.

The reason why an opt-out for all behavioral advertising is critical is that with the advent of cross-device tracking, previous identifiers that were not PII can become PII. The FTC cross-device tracking study detected non-PII

identifiers sent to the same third party services on different devices. But when "those devices share common attributes — such as the same local network and IP address — those services may be able to correlate user activity across devices" [25]. In that study 73 of 100 studied sites had privacy policies that reserved considerably broader rights to use and share "non personally identifiable information" like cookies and IP addresses. This same data "could be used for probabilistic cross-device correlation as well, by — for example — looking for devices that share IP addresses during certain periods of the day" [25]. This is also a reason why privacy policy specificity is useful. At the FTC seminar on cross-device tracking, one panelist argued that as datasets become more "easily cross-referenceable and aggregable, the distinction between personally identifiable information and non-personally identifiable information may diminish" [34].

While a company may claim to only transmit non-PII to third parties, the lines can get blurry. Especially since companies certainly have claim to need some single device tracking behavior to accomplish software engineering production changes for licensing, UX design, QA, etc. Companies will and should argue that some tracking is necessary for providing online services [30], which does have merit. However it is possible to still accomplish these goals if some percentage of consumer's opt-out of behavioral advertising as the company could anonymize that consumer data (but would have to clarify exactly how they do that in their privacy policy).

It is going to get difficult to regulate such tracking because so much of the data sharing will be on the back end, with first parties doing the cross-device tracking. Additionally it is technically difficult to comply with a consumer's full opt out depending on how data gathering is implemented by a first or third party. As a result, this thesis concurs with the opinion of computer security researcher Seda Gürses, who said a smart policy "will maybe limit itself to known players in the advertisement industry" [30] for full opt-out capability while smaller players would be expected to try in good faith to implement this to the best of their ability (and publicly admit they did so to be held liable). This thesis advocates specifically working with the top 10 advertising and tracking services that operate in both the web and mobile space as determined by the FTC. According to the study does by this thesis, the top 10 most prevalent ATS's are Arbor Technologies, LiveRamp, Tapad, Conversant, Signal, Neustar, AdNexus, AdAge, Bluekai, TraversedIP, and American List Counsel, Inc.

Collaboration with the DAA will be key because the DAA has the best working relationship with companies. Companies know that they should engage consumers in a way that will not cause them to lose trust in the marketplace [34] but are always skeptical of initiatives coming from an organization that can levy indictments against them. The FTC already has acknowledged the good work that the DAA is working towards in the cross-device tracking space stating that FTC commends their "self-regulatory efforts to improve transparency and choice in the cross-device tracking space. Both the NAI and DAA have taken steps to

keep up with evolving technologies and provide important guidance to their members and the public. Their work has improved the level of consumer protection in the marketplace" [9].

Companies will need to be mindful of the representations they make or risk violating the Section 5 authority of the FTC which prohibiting deception or unfairness in commerce "if they provide opt-outs that are unclear or deceptive, or that conflict with consumer expectations" [34]. Additionally the same warning would apply "to publishers who describe third-party opt-out programs in their privacy policies" [34].

Worried companies must be reassured that the FTC respects their right to serve and monetize advertisements. In the past, "blocking of ads and blocking of third-party trackers have been closely integrated, and seen as instances of the same problem" [23]. But this thesis concurs with an ad blocking paper entitled "The future of ad blocking: analytical framework and new techniques" by Narayanan et al. that advocates policy to separate the two sayings that "users might defend against [trackers] through anonymization techniques, faking cookies, etc. [and this would] diverge entirely from those involved in ad blocking" [23].

To distribute and inform advertising and tracking services of these policies the FTC, this thesis reorganized and compiled a list of publically available tracker domains [28] at bit.ly/CreepiesCrawliesAdTrackersList. The FTC could acquire the email addresses of the domain holders and contact them.

Engage in a robust education campaign

The FTC must begin a robust education campaign to talk to consumers and importantly, developers of mobile and web applications. Developers are rarely considered by policy making bodies as good targets for education because they are perceived to be complicit and knowledgeable on technical policy given their technical background. However while a lot of developers that integrate trackers from advertisers (such as using ad-libraries in developing apps) do so because they need money for getting their business of the ground, certainly "developers may also not be aware that they are doing so" [30]. Hence, it is important to "communicate consequences and best practices (and maybe also worst practices) in the industry to developers" [30].

The FTC should come up with a viral social media campaign as well as host workshops for concerned consumers to engage the public and educate them on cross-device tracking awareness. Some tips for consumers that the FTC has already written about include (and might be helpful for readers of this thesis):

-Use of a virtual private network (VPN) or Tor browser: this offers additional protection against linkability, though at a cost to performance (and in the case of a VPN, the cost of the service itself) [25]

-Resetting identifiers on mobile: iOS users can do this by following Settings > Privacy > Advertising > Reset Advertising Identifier. For Android, the path is Google settings > Ads > Reset advertising ID. This control works much like

deleting cookies in a browser — the device is harder to associate with past activity, but tracking can start anew using the new advertising identifier [38]

-Limit ad targeting on mobile devices: If you turn on this setting, apps are not permitted to use the advertising identifier to serve consumers targeted ads. For iOS, the controls are available through Settings > Privacy > Advertising > Limit Ad Tracking. For Android, Google Settings > Ads > Opt Out of Interest-Based Ads. Although this tool will limit the use of tracking data for targeting ads, companies may still be able to monitor your app usage for other purposes, such as research, measurement, and fraud prevention [38]

-Use tracker blocking software: consumers who wish to prevent or restrictively limit cross-device tracking can look into the use of tracker blocking software [25]

-Using “optout.aboutads.info/#/” a consumer can learn which third parties are tracking him or her and attempt to opt out of all tracking done by third parties work with the DAA

Though this thesis explores three particular policy options, there are several should be given thought by other readers. Other researchers have identified shortcomings in FTC reports and made suggestions such as the paper *Privacy leakage vs. protection measures: the growing disconnect* by Krishnamurthy et al. which offers options that a tracking blocker could enact and which options would stop expected, known or potential PII leakage [5] as listed in Figure 8.

Table 2: Effectiveness of Protection Measures for a) Expected; b) Known; and c) Potential Leakage and Linkage Scenarios

		Protection Measure							
	Leakage/Linkage Scenario	block	block-hidden	nocook	no3rdcook	nojs	referer	anon	opt-out
a) Expected	User visit	X							
	Hidden third party		X						
	3rd-party tracking linkage	X		X	X				X
	1st-party tracking linkage			X		X			
b) Known	Leakage via Referer	X					X		
	Leakage via cookies		X	X					
	Leakage via JS	X				X			
c) Potential	Linkage via IP addr	X						X	
	Linkage via Flash cookies	X							
	Linkage via fingerprint	X				X			
	Linkage via GUID	X					X		
	Linkage w/ Other Sources								

Figure 8: Protection measures that a tracking blocker could offer as described by Krishnamurthy et al.

A cross-device tracking blocker might work since “even though publishers increasingly deploy scripts to detect and disable ad blocking, ad blockers run at a higher privilege level than such scripts, and hence have the upper hand in” [23] the back and forth over consumer privacy and security. However a government agency is unlikely to develop such a blocker so it would have to be a private sector solution.

Individual users cannot really do much since the ecosystem is moving towards more tracking and more authentication with software as a service solutions. Seda Gürses recommends consumers and researchers read “recent papers that try to obfuscate against cross-device/app tracking and or block third party ads/libraries” [30] to think of potential solutions. Specifically for mobile tracking, the FCC has been thinking about whether there “are there any ways in

which [their] existing and proposed notice requirements can or should be tailored to the unique characteristics of mobile services and smaller screens [21]?"

A last party that should not be neglected in this conversation is the role of first party sites in safeguarding consumer privacy, a segment that Krishnamurthy et al. have pointed out was left out of the conversation and a "a key failure of the [2010 consumer privacy] FTC report" [5]. First party sites (the Fitbit's and Pandora's) should also be held responsible for any data they knowingly or unknowingly transmit to third party trackers.

If the policy options laid out by this thesis are enacted, negative cross-device tracking consequences can be slowed and minimized. However the scrutiny must be ongoing as cross-device tracking can be performed in the future on any current or future data collected. It is possible that limited third-party cross-device tracking is happening today, "though any retained data could be used for ex post cross-device correlation in the future unless there are contractual prohibitions on this usage [25]."

Conclusion

Cross-device tracking has many benefits. It allows for "seamless, consistent consumer experiences across devices and better techniques for protecting consumers from fraud. It also allows for improved ad efficiency, reduced ad fatigue, and better monetization practices" [34]. However

cross-device tracking raises certain privacy concerns. FTC Chairwoman Edith Ramirez said it best when she said that “cross-device tracking blurs the line between aspects of consumers’ lives that they may intend to keep separate” [34].

The cross-device tracking space will continue to grow as companies are created in this fledgling industry. Existing companies are also eying cross-device tracking as a revenue stream to expand into. 25 companies just from this study alone have the potential to be currently engaging in cross-device tracking. Integral Ad Science for example has the method on their roadmap for cross channel verifications to tie the ad campaigns together. This would really not affect the consumer because it would be to get metrics for what ad impressions performs best (and Integral Ad Science can only get data from inside an ad, not consumer login/browsing behavior) but is still a move into the space [22-skovron].

Better privacy policies, single opt-out policies and consumer/developer education are three key ways to reduce negative impacts of cross-device tracking. Companies, and not consumers, seem to benefit most from cross-device tracking [34] a practice which is just creepy. As the landscape evolves consumers will have more of a say in how and what is being collected about them. Companies have to pay more attention to privacy sites and regulators due to loss of brand value associated with not being privacy conscious [5]. As legendary security expert Bruce Schneier said, “If more people

had a security mindset, services that compromise privacy wouldn't have such a sizable market share -- and Facebook would be totally different."[11]

Remember the plea of the self-regulatory body of advertising and tracking companies: "let's be sure we keep true to our principles of enhanced transparency and consumer control. The reward is better consumer engagement and confidence through a responsible internet-based advertising ecosystem with meaningful accountability." [35]

Bibliography

1. Englehardt, S., & Narayanan, A. (2016). Online Tracking. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security - CCS'16*, (1), 1388–1401. http://doi.org/10.1145/2976749.2978313
2. Acker, S. Van, Hausknecht, D., & Sabelfeld, A. (2017). Measuring Login Webpage Security, 1–8. https://doi.org/10.1145/3019612.3019798
3. Acker, S. Van, Hausknecht, D., & Sabelfeld, A. (2017). Measuring Login Webpage Security, 1–8. https://doi.org/10.1145/3019612.3019798
4. Starov, O., Gill, P., & Nikiforakis, N. (2016). Are You Sure You Want to Contact Us? Quantifying the Leakage of PII via Website Contact Forms. Proceedings on Privacy Enhancing Technologies, 2016(1), 20–33. https://doi.org/10.1515/popets-2015-0028
5. Krishnamurthy, B., Naryshkin, K., & Wills, C. E. (2011). Privacy leakage vs. protection measures: the growing disconnect. *Web 2.0 Security and Privacy Workshop.*, 1–10.
6. Federal Trade Commision (FTC). (2012). Protecting Consumer in an Era of Rapid Change: Recommendations for businesses and policymakers. *Federal Trade Commision*, (March), 1–112. Retrieved from https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf
7. Federal Trade Commission. "Re: Comments for November 2015 Workshop on Cross-Device Tracking" Center for Democracy & Technology, pp 1-11. Washington, DC. 2015. https://cdt.org/files/2015/10/10.16.15-CDT-Cross-Device-Comments.pdf
8. Justin Brookman. "Cross-device tracking, an FTC Workshop." Federal Trade Commission. Slide 1-41. Washington, DC. 2015. https://docs.google.com/presentation/d/1_wKwr7l_rhILTuSnqAM4l_NfsRA5atAKIm0_fJnMp8k/edit#slide=id.p4

9. FTC Staff Report. (2017). Cross-Device Tracking (January). https://www.ftc.gov/system/files/documents/reports/cross-device-tracking-federal-trade-commission-staff-report-january-2017/ftc_cross-device_tracking_report_1-23-17.pdf
10. Michael Whitener, (2015) Cookies Are So Yesterday; Cross-Device Tracking Is In—Some Tips. https://iapp.org/news/a/cookies-are-so-yesterday-cross-device-tracking-is-insome-tips/
11. Schneier, Bruce (2008). The Security Mindset. https://www.schneier.com/blog/archives/2008/03/the_security_mi_1.html
12. Acar, G., Eubank, C., Englehardt, S., Juarez, M., Narayanan, A., & Diaz, C. (2014). The Web Never Forgets: Persistent Tracking Mechanisms in the Wild. *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security - CCS '14*, 674–689. https://doi.org/10.1145/2660267.2660347
13. Englehardt, S., Reisman, D., Eubank, C., Zimmerman, P., Mayer, J., Narayanan, A., & Felten, E. W. (2015). Cookies That Give You Away. *Proceedings of the 24th International Conference on World Wide Web - WWW '15*, 289–299. https://doi.org/10.1145/2736277.2741679
14. Chanchary, F., & Chiasson, S. (2015). User Perceptions of Sharing, Advertising, and Tracking. *Symposium on Usable Privacy and Security (SOUPS) 2015, July 22–24*, 53–67.
15. Roesner, F., Kohno, T., & Wetherall, D. (2012). Detecting and defending against third-party tracking on the web. *Proc. of the USENIX Conference on Networked Systems Design and Implementation (NSDI)*, (Nsdi), 12.
16. Aleyasen, A., Starov, O., Au, A. P., Schiffman, A., & Shrager, J. (2015). On the privacy practices of just plain sites. Proceedings of the 14th ACM Workshop on Privacy in the Electronic Society, 1–10. http://doi.org/10.1145/2808138.2808140
17. Huber, Markus, Martin Mulazzani, Sebastian Schrittwieser, and Edgar R. Weippl. "Large-scale Evaluation of Social Apps." *AppInspect* (n.d.): 1-18. ACM COSN, Boston, 8 Oct. 2013. Web .https://www.sba-research.org/wp-content/uploads/publications/AppInspect_peprint.pdf
18. Shubro Saha, A large-scale, dynamic analysis of user privacy in Android applications, Princeton Senior Thesis, Department of Computer Science. 2015. pp. 1-20
19. Vallina-Rodriguez, Narseo, Srikanth Sundaresan, Abbas Razaghpanah, Rishab Nithyanand, Mark Allman, Christian Kreibich, and Phillipa Gill. "Towards Understanding the Mobile Advertising and Tracking Ecosystem." *Tracking the Trackers* (n.d.): 1-6, 26 Oct. 2016. Web. https://arxiv.org/pdf/1609.07190.pdf.
20. Narayanan, A., & Reisman, D. (2017). The Princeton Web Transparency and Accountability Project, 1–24.
21. Wheeler, Clyburn, Rosenworcel, Pai, and O'Rielly. "Customer Proprietary Network Information." *Before the Federal Communications Commission Washington, D.C.* (n.d.): 102-03. 1 Apr. 2016. Web. https://apps.fcc.gov/edocs_public/attachmatch/FCC-16-39A1.pdf.
22. Skovron, John. (2016). Personal Interview Phone Call.
23. Storey, G., Reisman, D., Mayer, J., & Narayanan, A. (2017). The Future of Ad Blocking: An Analytical Framework and New Techniques. Retrieved from http://randomwalker.info/publications/ad-blocking-framework-techniques.pdf

24. Mayer, Jonathan R. ""Any Person... a Pamphleteer" Internet Anonymity in the Age of Web 2.0." (n.d.): 1-103. Jonathan R. Mayer, 7 Apr. 2009. Web. https://jonathanmayer.org/papers_data/thesis09.pdf.
25. Brookman,, Justin, Phoebe Rouge, Aaron Alva, and Christina Yeung. "Cross-Device Tracking: Measurement and Disclosures." *Proceedings on Privacy Enhancing Technologies* (n.d.): 134-49. 01 Dec. 2017. Web. https://petsymposium.org/2017/papers/issue2/paper29-2017-2-source.pdf.
26. Cao, Yinzhi, Song Li, and Erik Wijman. "(Cross-)Browser Fingerprinting via OS and Hardware Level Features." (n.d.): 1-15. U.S. National Science Foundation, 27 Feb. 2017. Web. http://yinzhicao.org/TrackingFree/crossbrowsertracking_NDSS17.pdf.
27. Ren, Jingjing, Ashwin Rao, Martina Lindorfer, Arnaud Legout, and David Choffnes. "ReCon: Revealing and Controlling PII Leaks in Mobile Network Traffic." (n.d.): 1-18. Data Transparency Lab, 19 Aug. 2016. Web. https://arxiv.org/pdf/1507.00255.pdf.
28. Rao, Ashwin, Arash Molavi Kakhk, Abbas Razaghpanah, Amy Tang, Shen Wang, Justine Sherry, Phillipa Gill, Arvind Krishnamurthy, Arnaud Legout, Alan Mislove, and David Choffnes. "Using the Middle to Meddle with Mobile." (n.d.): 1-14. 4 Dec. 2013. Web. http://david.choffnes.com/pubs/meddle-main.pdf.
29. Gürses, Seda, and Joris Van Hoboken. "Privacy After the Agile Turn." *The Cambridge Handbook of Consumer Privacy* (n.d.): 1-29. Selinger Et Al, 2017. Web. https://drive.google.com/file/d/0B0NjQdX1kw4sSU1odTA5eG1EV0k/view.
30. Gurses, Seda (2017). Email Interview.
31. Mayer, John (2011). TRACKING THE TRACKERS: WHERE EVERYBODY KNOWS YOUR USERNAME. http://cyberlaw.stanford.edu/blog/2011/10/tracking-trackers-where-everybody-knows-your-username
32. Diaz-Morales, R. (2016). Cross-Device Tracking: Matching Devices and Cookies. *Proceedings - 15th IEEE International Conference on Data Mining Workshop, ICDMW 2015*, 1699–1704. https://doi.org/10.1109/ICDMW.2015.244
33. Digital Advertising Alliance (2015). Application of the Self-Regulatory Principles of Transparency and Control to Data Used Across Devices. http://www.aboutads.info/sites/default/files/DAA_Cross-Device_Guidance-Final.pdf
34. Friel, Alan & Goldberg, Daniel (2015). The FTC and DAA Set Their Sights on Cross-Device Tracking. https://www.dataprivacymonitor.com/behavioral-advertising/the-ftc-and-daa-set-their-sights-on-cross-device-tracking/
35. Digital Advertising Alliance (2017). Cross Device Guidance DAA Principles Enforcement Beings Feb 1 http://digitaladvertisingalliance.org/blog/cross-device-guidance-daa-principles-enforcement-begins-feb-1-2017/
36. Ohm, Paul (2012). Don't Build A Database of Ruin. https://hbr.org/2012/08/dont-build-a-database-of-ruin
37. Brodkin, John (2017). FCC to halt rule that protects your private data from security breaches.

https://arstechnica.com/tech-policy/2017/02/isps-wont-have-to-follow-new-rule-that-protects-your-data-from-theft/
38. Federal Trade Commission (2017). Controlling Online Tracking. https://www.consumer.ftc.gov/articles/0042-online-tracking#Controlling_Online_Tracking
39. Narayanan, Arvind (2011). The Linkability of Usernames. https://33bits.org/2011/02/16/usernames-linkability-uber-profiles/
40. Python Script by Max Greenwald - Hashes.py
41. Bash Script by Max Greenwald - Bash Script
42. Python Script by Max Greenwald - csvMobileData.py
43. OpenWPM - an opensource framework
44. Python Script by Max Greenwald - ExtractPII.py

Appendices

Appendix A: Code

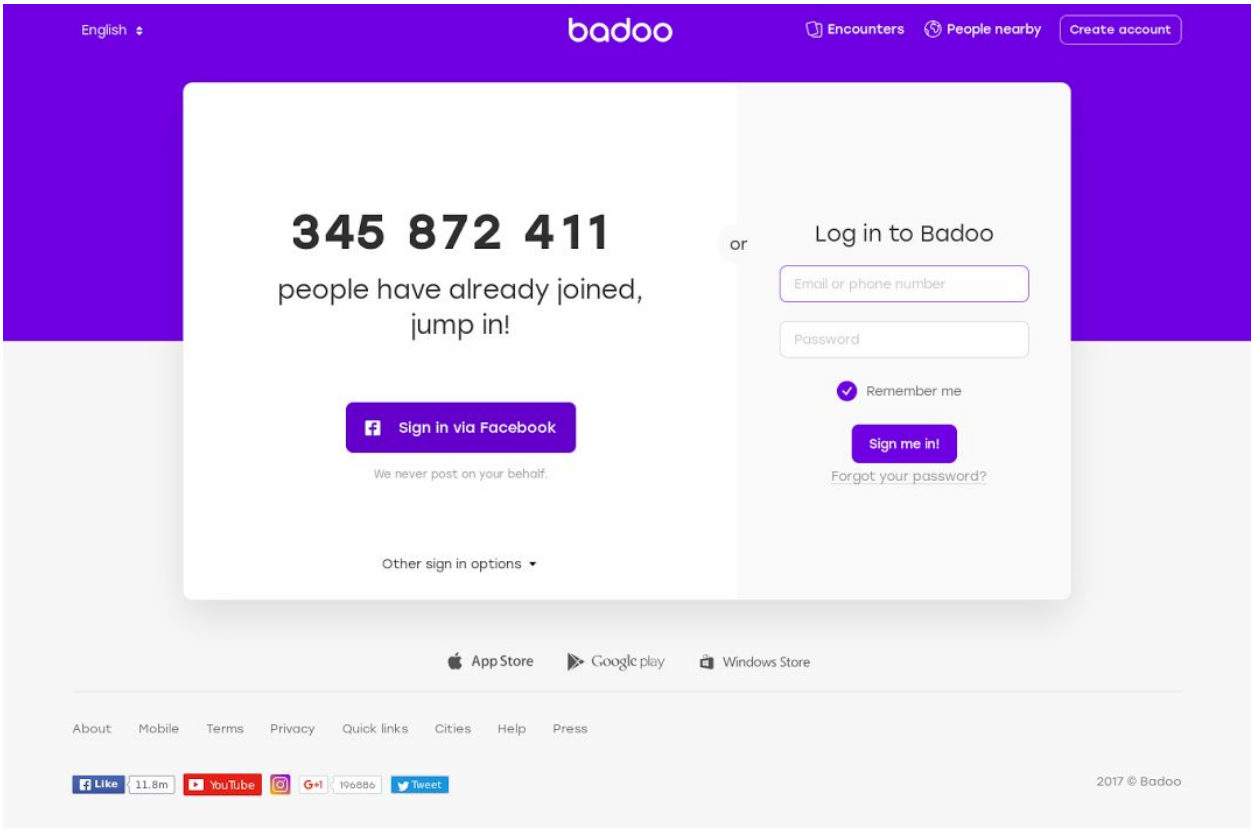
All mobile code except for some of the cleaning/comparing scripts are available on my github at https://github.com/MaxGreenwald/Cross-Device-Tracking. The Web Code will be available through OpenWPM under Facebook Login by late 2017.

Appendix B: Future Research Opportunities

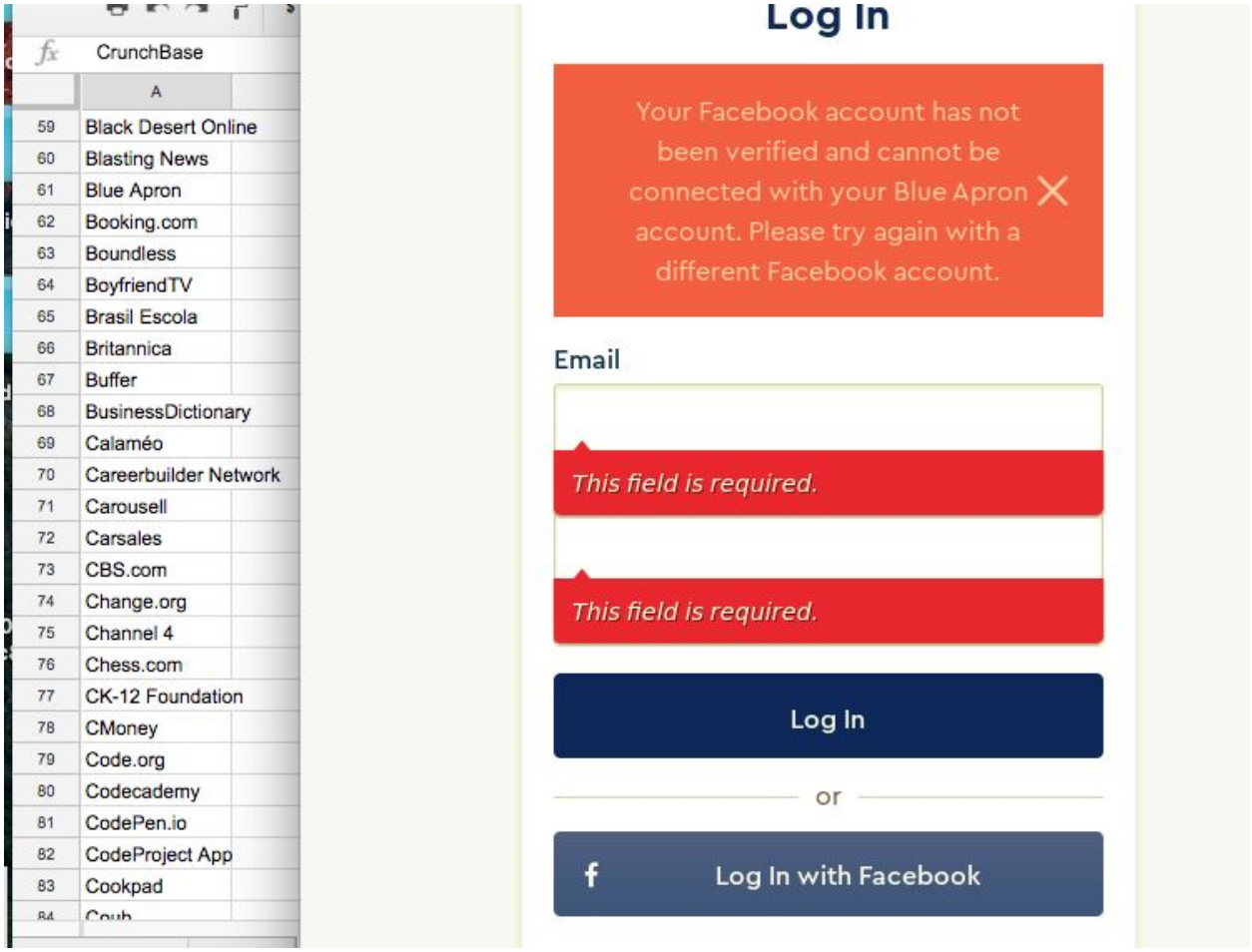
1. Bring privacy policies into the conversation: perhaps compare privacy policies to sites permitting/implementing cross drive tracking to see who is in violation of their own privacy policy. See a study on financial institution privacy practices
2. Automated mobile analysis to get a larger mobile study conducted
3. Survey Android apps for cross device tracking
4. Get in touch with 10 third party tracking companies and understand more about their motivations and intentions

Appendix C: Facebook Login Crawler

Creating a successful login with Facebook crawler is difficult. There are a lot of necessary steps one needs to take to verify that a site is successfully logged into. First the crawl has to have reached a facebook.com/login.php page and entered some fake credentials (I was able to do this for 865 sites of the 7561 I crawled). From there you should be able to query the Facebook API to make sure that you're logged in but unfortunately many sites don't include the Facebook Object post login. Only 88 of those 865 sites I reached a Facebook login page for were were "fb_verified" meaning we pinged the Facebook API and it confirmed we were currently logged in. Confusing though because for some of the "fb_verified" sites (such as badoo below) shows that we have not logged in yet so you cannot fully trust even pinging the Facebook API.

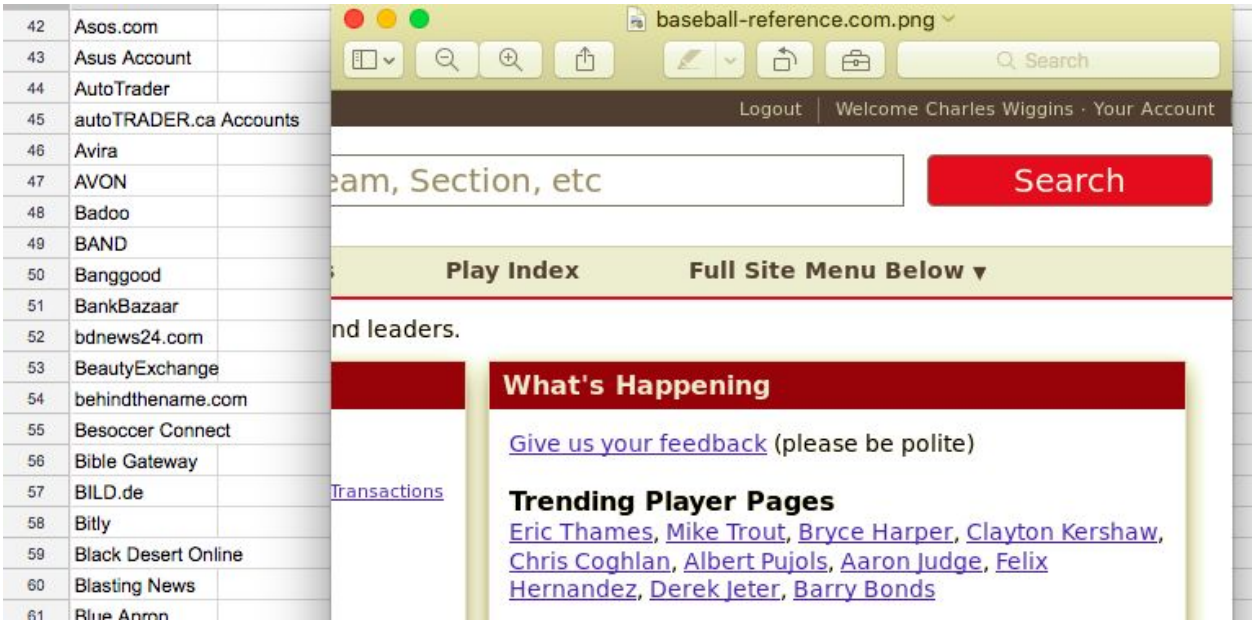


One might think that the list of Connected Apps Through Facebook would be the sites that are officially connected and successfully logged in to that site however only 25% of those apps had screenshots that showed I was not logged in like Blue Apron below.



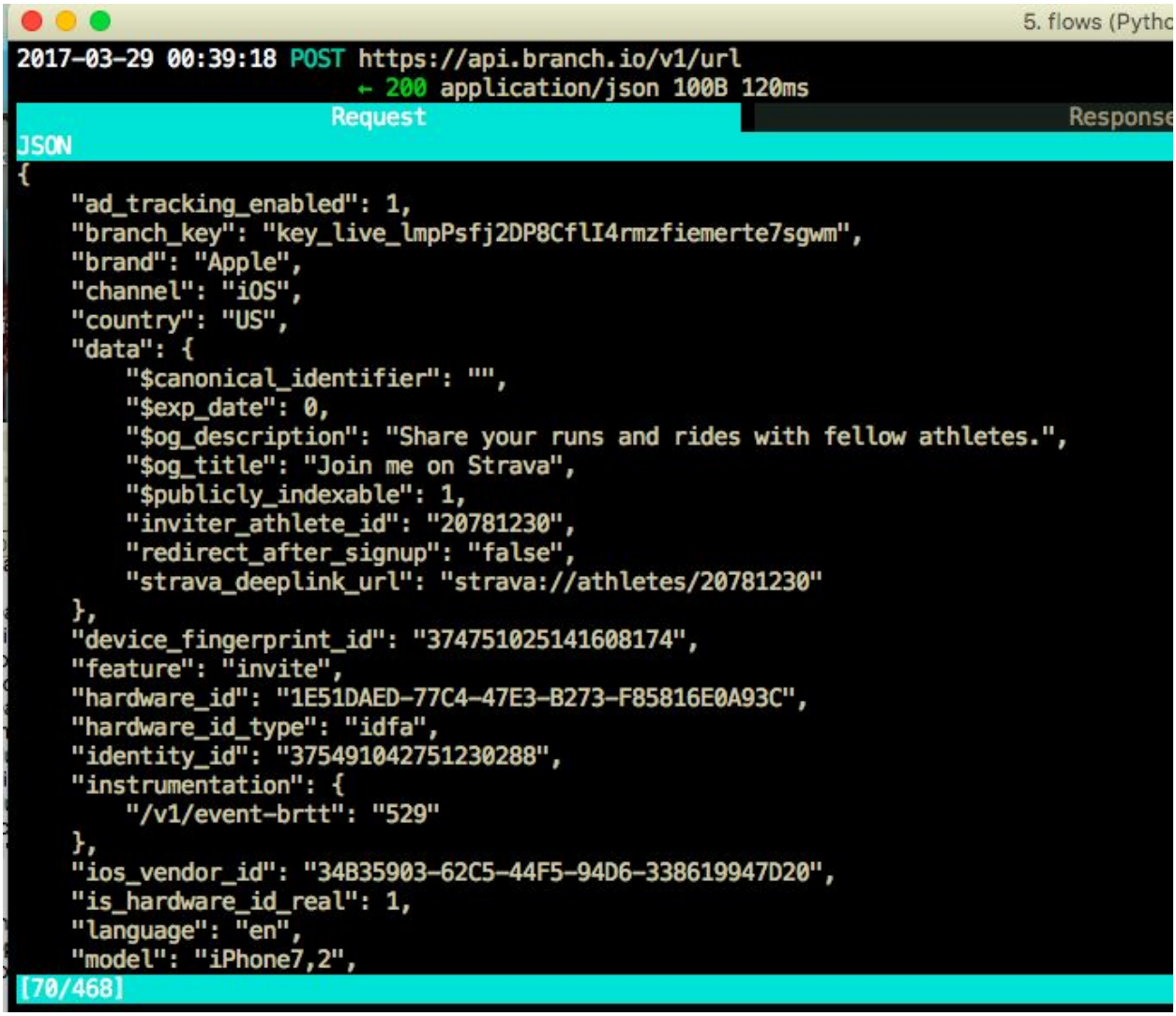
Blue apron shows connected through facebook but not logged in in the screenshot

Baseball-reference has the opposite problem where the screenshot indicates that we are successfully logged in but the app is not connected through the Facebook Connected Apps display.



Appendix D: Mitmproxy Screenshots

Here are some examples of how the unfiltered data looked before the Python scripts pulled out the plaintext or hashed PII


5. flows (Pytho
2017-03-29 00:39:18 POST https://api.branch.io/v1/url
← 200 application/json 100B 120ms
Request
Response
JSON
{
"ad_tracking_enabled": 1,
"branch_key": "key_live_lmpPsfj2DP8CflI4rmzfiemerte7sgwm",
"brand": "Apple",
"channel": "iOS",
"country": "US",
"data": {
"$canonical_identifier": "",
"$exp_date": 0,
"$og_description": "Share your runs and rides with fellow athletes.",
"$og_title": "Join me on Strava",
"$publicly_indexable": 1,
"inviter_athlete_id": "20781230",
"redirect_after_signup": "false",
"strava_deeplink_url": "strava://athletes/20781230"
},
"device_fingerprint_id": "374751025141608174",
"feature": "invite",
"hardware_id": "1E51DAED-77C4-47E3-B273-F85816E0A93C",
"hardware_id_type": "idfa",
"identity_id": "375491042751230288",
"instrumentation": {
"/v1/event-brtt": "529"
},
"ios_vendor_id": "34B35903-62C5-44F5-94D6-338619947D20",
"is_hardware_id_real": 1,
"language": "en",
"model": "iPhone7,2",
[70/468]

```
      age=610&ns_ap_fg=2&ns_ap_ft=3107&ns_ap_dft=3107&ns_ap_bt=606&ns_ap_dbt=606&ns_ap_dit=5433
      11945&ns_ap_as=1&ns_ap_das=32135&ns_ap_it=543311945&ns_ap_us=1&ns_ap_dus=3031531&ns_ap_ut
      =120000&ns_ap_lang=en-US&ns_ts=1490758514933
      ← 200 image/gif 43B 86ms
>> HEAD http://i'mgettinghotandbotheredwithmyhand/
       ← ProxyError(NetLibError('Error connecting to "i\'mgettinghotandbotheredwithmyhand":
         [Errno 8] nodename nor servname provided, or not known',),)
  GET https://graph.facebook.com/v2.2/267470608481?fields=name,supports_implicit_sdk_logging,gd
      pv4_nux_enabled,gdpv4_nux_content,ios_dialog_configs,app_events_feature_bitmask&sdk=ios&f
      ormat=json
      ← 200 application/json 298B 170ms
  GET https://www.gstatic.com/identity/sdk/config/ios/v2.3.0?clientID=this%20plist%20needs%20to
      %20be%20in%20the%20project%20file%20in%20order%20for%20the%20config%20to%20init.%20All%20
      client%20id%20can%20be%20found%20on%20SuServiceGoogle.m&bundleID=stumbleupon
```

2. flows (Python)

```
      ← 302 text/html 361B 996ms
  GET https://s3-media1.fl.yelpcdn.com/bphoto/BUZXbkOQk6tqw0ZWNkg_fg/l.jpg
      ← 200 image/jpeg 36.69kB 949ms
  GET https://s3-media2.fl.yelpcdn.com/bphoto/TEpxtghVqPlKdlI5Zg2T7w/168s.jpg
      ← 200 image/jpeg 5.41kB 1.10s
  GET https://www.google-analytics.com/r/collect?v=1&_v=j52&a=77947504&t=pagevi
      ew&_s=1&dl=https%3A%2F%2Fwww.yelp.com%2Fjersey&dp=%2Fhome%2Fjersey&ul=en-
      us&de=UTF-8&dt=Jersey%20City%20Restaurants%2C%20Dentists%2C%20Bars%2C%20B
      eauty%20Salons%2C%20Doctors%20-%20Yelp&sd=24-bit&sr=1440x900&vp=1109x684&
      je=0&fl=21.0%20r0&_u=QICAAAABI~&jid=1131912721&gjid=1071739614&cid=49297C
      1DAC9790FB&tid=UA-30501-24&_r=1&cd1=anon&cd81=status_quo&cd53=1&cd189=sta
      tus_quo&cd108=status_quo&cd185=status_quo&cd183=six_pack&cd34=%2Fjersey&c
      d120=enabled_logged_in_and_logged_out&cd64=none&cd138=1&cd27=False&z=7950
      4185
      ← 302 text/html 360B 1.41s
  GET https://s3-media2.fl.yelpcdn.com/bphoto/AMTBLnuF4uoy3prUSwQGYA/300s.jpg
      ← 200 image/jpeg 13.82kB 972ms
  GET https://s3-media4.fl.yelpcdn.com/bphoto/Hu6-pYnhpM5Q_ZkC0e04uQ/168s.jpg
      ← 200 image/jpeg 8.51kB 841ms
  GET https://s3-media2.fl.yelpcdn.com/photo/DvEiSoaS9drqHtoYU0B_lA/30s.jpg
      ← 200 image/jpeg 870B 794ms
>> GET https://s3-media4.fl.yelpcdn.com/photo/HqO-lKyhwBVrsf22Bsar6A/30s.jpg
      ← 200 image/jpeg 1003B 768ms
[67/625]                                                    ?:help [*:8080]
```