Geolocation Tracking and Privacy Issues Associated with the Uber Mobile Application

Dr. Darren R. Hayes
Pace University
dhayes@pace.edu
Sapienza Università di Roma
darren.hayes@uniroma1.it

Christopher Snow
csnow@pace.edu

Saleh Altuwayjiri
sa07549n@pace.edu

Pace University
New York, NY

Abstract
This research examined the Uber mobile application and determined that the company utilizes very precise and potentially invasive geolocation tracking techniques. Our experiments indicate that Uber tracks the location of its users, after the conclusion of a ride, for longer than its official privacy policy would indicate. The most interesting finding is that geolocation tracking is performed by the Uber app even when the user does not take an Uber ride. Additionally, its geolocation tracking of users using competing services may be disconcerting for some consumers. While our findings may be a privacy concern for some consumers, the Uber app also has tremendous potential for digital forensics investigators.

Keywords: Uber, mobile apps, mobile forensics, geolocation, privacy

1. INTRODUCTION
Uber is a service that enables drivers to act as flexible contractors and provide transportation services that compete with traditional taxi services. Consumers, using the Uber mobile app, can search for a car service in their area. The benefit to the consumer is that they are visually provided with the mapped location of Uber cars in their locality and be provided with an upfront quote for a specific journey (or “ride”). Uber operates in 570 cities worldwide. In recent times, Uber has received negative press about its geolocation tracking of users and raised a number of concerns regarding its privacy policies and potentially invasive data collection practices. The research, herein, sought to identify the type of personally identifiable information (PII) collected by the company. More precisely, this research identifies what geolocation data the Uber mobile application (app) collects on the user.

In April 2017, the New York Times published a story that documented a meeting, at Apple headquarters, in 2015, between Travis Kalanick, CEO of Uber, and Tim Cook, CEO of Apple (Issac, 2017). The article alleges that Mr. Cook scolded Mr. Kalanick for identifying and tagging iPhones after the Uber app had been uninstalled.
or the device had been wiped. Apparently, this type of user identifying coding violated the Apple developer terms of service agreement (Newman, 2017).

A recent article in The New York Times detailed how Unroll.me, which purported to purge your device’s mail Inbox of annoying advertising messages, was actually being used to spy on competitors (Isaac & Lohr, 2017). The article documented how Unroll.me would scan a user’s Inbox, identify if there were service receipts, from competing companies like Lyft, and then sell that information to Lyft’s competitor - Uber.

2. BACKGROUND

Since the introduction of iOS 5, Apple has been limiting app developer access to the iPhone’s UDID (Unique Device Identifier) (Schonfeld, 2011). A notice from Apple stated, “Starting May 1, the App Store will no longer accept new apps or app updates that access UDIDs. Please update your apps and servers to associate users with the Vendor or Advertising identifiers introduced in iOS 6” (Panzarino, M.). Apple now prefers that app developers utilize the official Apple advertising platform to track app users. Based on Apple’s Advertising and Privacy policy, it appears that Apple does collect user data and then subsequently shares it with third-parties (Apple, 2017). Nevertheless, developers can obtain extensive information about an app user through the integration of the UIDevice object. The UIDevice object can be used by an app developer to determine the assigned name of the device, device model and iOS version, orientation (orientation property) of the device, battery charge (batteryState property) and distance of the device to the user (proximityState property) (Apple). Moreover, developers can integrate code, during app development, for third-party analytics into their app. These third-party companies include Localytics, mixpanel, UXCam and fabric. Companies like Apptopia provide app developers with extensive, nay invasive, analytics on competitor apps.

The use of the user UDID has not always been used for nefarious purposes, however; the UDID was often used to identify if the user of an app was a legitimate user and could block a user’s access if an account was compromised or potentially stolen. Fingerprinting is yet another methodology, used by third-parties, to uniquely identify users, based on application configuration. Fingerprinting is best known for identifying online users based on user settings from their browser, which may include user cookies and browser plug-ins. The Electronic Frontier Foundation (EFF) created a project known as Panopticlick (panopticlick.eff.org) to raise awareness about how your browser is used by advertisers, and others, to identify and track you on the Web. The EFF announced that 84% of online users can be uniquely identified by their browser (Budington, 2015).

According to Uber’s USER PRIVACY STATEMENT, there are two categories of information collected about users: (a) Information You Provide to Us, which can include name, email, phone number, postal address, profile picture, payment method, and (b) Information We Collect Through Your Use of Our Services, which can include location information, contacts, transactions, usage and preference, device information, call and SMS data and log information (Uber, 2015). Of particular interest is the Device Information (hardware model, operating system and version, software and file names and versions, preferred language, unique device identifier, advertising identifiers, serial number, device motion information, and mobile network information). In terms of Location Information, Uber is not specific about the extent to which the user’s location is being tracked but states that they “may also collect the precise location of your device when the app is running in the foreground or background” (Uber, 2015). Uber provides more detailed information about the use of Location Services on its Website under iOS App Permissions (Uber).

What is most interesting, for the purposes of our research, is that during our installation of the Uber app, a dialog box appears and states that “Uber collects your location (i) when the app is open and (ii) from the time of the trip request through five minutes after the trip ends” (see Figure 1). Uber states in their FAQ that the reasoning behind this data collection is to “improve pickups, drop-offs, customer service, and to enhance safety” (Uber). However, users reported seeing the Uber app using location services weeks after the app was used, past the 5 minutes they claim. Uber responded to these reports blaming Apple’s iOS Maps extension that Uber uses to serve regional maps to their customers (Perez, 2016).

Perhaps unsurprisingly, Uber has invested heavily in data science to retain its competitive advantage, as evidenced by its aggressive recruitment of data scientists (Wilde, 2015). We
also know that Uber extensively uses a telematics pilot program, called Autohawk, to identify the location of its drivers and perform diagnostic testing on the vehicle to ensure passenger safety (Wiszniowski, 2016). In fact, Uber provides geolocation information, provided by its data visualization team, on its website at eng.uber.com/data-viz-intel. Uber integrates both Fabric and Localytics in its mobile app. Fabric provides companies, like Uber, with real-time information about the health of their app. These analytics include application crash analytics.

3. RELATED WORK

The research group at Pace University has previously published research on how geolocation information from mobile apps could be used by governments to track users in an article entitled “Leakage of Geolocation Data by Mobile Ad Networks” (Snow, Hayes, Dwyer, 2016) and another article entitled “Mobile Ad Networks and Security Issues Regarding Geolocation Data” (Snow, Hayes, Dwyer, 2016). More specifically, this research focused on the claims of the whistleblower, Edward Snowden, that mobile apps, like Angry Birds, could be used to profile individuals and track their whereabouts through communications with mobile advertisements.

EURCOM’s research paper, entitled “Taming the Android AppStore: Lightweight Characterization of Android Application”, examined the network connections established by “legitimate” apps (Vigneri, Chandrashekar, Pefkianakis, & Heen, 2015). Ultimately, their research determined that the 2,146 applications that they examined connected to approximately 250,000 distinct URLs; many of these server connections were to domains known to maintain malware.

A recent paper published by IEEE titled “Location Privacy Breach: Apps Are Watching You in Background” (Liu, D., Gao X., & Wang, H.) examined the correlation between the frequency of locational requests and potential security risks associated with user data leakage. The study concluded that shorter intervals between each geolocation request from a mobile app correlate to a higher privacy risk. Therefore, mobile applications that constantly update with a user’s location, such as a car service are at high risk of leaking a user’s location because of their high request rate.

It is also known that mobile users are not generally aware when an app is accessing personal data in the background, as noted in the research study “Android permissions remystified: a field study on contextual integrity” published by USENIX Security (Wijesekera P., Baokar A., Hosseini A., Egelman S., Wagner D., and Beznosov K.). The 2015 study analyzed how 36 participants used their mobile applications and concluded that 80% of participants studied noted that at least one permission request, made by a mobile application, as inappropriate. These permission requests were approved when the user accepted the end-user agreement during installation. The participants were unaware about exactly what they were allowing the mobile application to access. Thus, a majority of the participants experienced an invasion of (locational) privacy from the mobile applications they used.

4. EXPERIMENTAL RESEARCH

Our methodology, for this research, involved performing both a static and dynamic analysis of the Uber mobile application for iOS (iPhone) and for Android. The static analysis primarily involved an analysis of the SQLite database associated with the app, utilizing mobile forensic tools on an iPhone. Additionally, our static analysis included reverse engineering the Android application package (APK) file. The purpose of the latter was to review the code to identify the application’s manifest. This app manifest would include permissions on the user device, which could include access to contacts, user location (based on GPS, cell sites or local access points), access to device hardware, like the camera or microphone and identifying information about the user. The dynamic analysis focused on how the application behaved during execution. The latter would use a series of DNS analytical tools to identify domain and network connections.

Google Maps

The Uber app for Android utilizes the Google Maps API (application programming interface) (see Figure 2) for locating Uber drivers in the vicinity of the consumer. We also confirmed this through our dynamic analysis of PCAPs from the Uber app (see Figure 3).

APK Analysis

An examination of the Uber APK file will provide a list of permissions requested by the application. A review of the APK manifest will quickly identify location permissions requested
of the user. A search for the “Uber APK file” will quickly identify where the application can be downloaded from the Web. Once downloaded there are a number of applications that can be used to review the code and manifest for the APK. One tool for reviewing APK developer code is dex2jar (dex compiler), which can be downloaded from SourceForge. Yet another application for viewing the APK is FileViewer Plus. For our research we chose to use an online APK decompiler application, which is available from www.javadecompilers.com/apk. The rationale for selecting this tool to decompile the APK is that no download is required and the APK file can simply be uploaded on the fly. An analysis of the APK manifest reveals the following permissions related to location:

android.permission.ACCESS_COARSE_LOCATION
android.permission.ACCESS_FINE_LOCATION

There is nothing unusual about these location permissions being requested by the Uber app. The location of a user device can be based on GPS, proximity to a cellular tower (cell site) or WiFi. ACCESS_COARSE_LOCATION is a permission that enables the app to access the approximate location of the user device, which is based on NETWORK_PROVIDER (cell sites). ACCESS_FINE_LOCATION enables the app to determine the location of the user device based on NETWORK_PROVIDER and GPS (GPS_PROVIDER).

Using BlackLight (BlackBag Tech) and MPE+ (AccessData) forensics tools, we can determine that the data associated with the Uber app can be found in the following application files:

Library>Application
Support>com.ubercab.UberClient\xxxx.idb

Library>Application
Support>com.ubercab.UberClient\Cache.db

Data>data>com.ubercab\files\ rider\0000xx.idb

Analysis Background
The data below was compiled using two analyst programs:
- Debookee by iwaxx
- BlackLight by BlackBag

The initial experiments were run from an iPhone 7 Plus device running iOS 10.2.1. The Uber app was downloaded for the first time on this new iPhone. The user account on the Uber app was new and no ride history existed. Actions that were performed include searching for a nearby Uber car, mapping out a pick-up point and destination address to determine ETA (estimated time of arrival) and price in addition to adding a PayPal payment method.

Man-in-the-Middle Analysis
The highlighted lines in Figure 4 are HTTPS requests made by the Uber app. These lines were captured using Debookee, a LAN scanning program by iwaxx; the tool can capture traffic on a targeted device. Each request uses the SSL, under the HTTPS protocol, as seen in the “method” column. The HTTPS protocol ensures that data being pushed from the user, like location, to the Uber servers is encrypted and therefore illegible to bad actors listening to the connection. One can infer that cn-geo1.uber.com is a domain used specifically for transferring information about a user’s location and nearby Uber vehicles.

Location Analysis with BlackLight
Using BlackLight, we were able to find an interesting file called “eyeball”, as seen in Figure 5.

```
```

The above string is a sampling of many similar strings that were found in a file labeled “eyeball” under the following folder pathway in the UberClient application folder:

```
Library>Application
Support>PersistentStorage>BootstrapStore>RealmApp.StreamModelKey
```

It can be inferred that “eyeball” refers to the radius circle of a user’s location to capture nearby Uber vehicles. This “etaString” operation is repeated throughout the “eyeball” storage file for multiple vehicles in the area. The string explains an Estimated Time of Arrival (ETA) of 5
minutes and provides an array of latitude and longitude numbers connected to a “vehicle path” followed by a string of numbers and letters that can be concluded as a serial number for the specific vehicle. When the various longitude and latitude coordinates in the array are mapped out, the points display a route that this specific nearby Uber vehicle was following. Figure 6 shows the latitude and longitude coordinates mapped out on Google Maps. It can be inferred that each of these coordinates was a ping from the nearby Uber vehicle to the app in order to update the driver’s position on the user’s screen.

The most important string from this same “eyeball” storage file came at the very end. The following string shows a “reverseGeocode” operation that was able to identify the exact address of where the user was located.

```
"reverseGeocode":{{"latitude":40.7102306652868, "components":{{"long_name":"163 William Street","short_name":"163 William Street"},"types":[]},
{"long_name":"Lower Manhattan","short_name":"Lower Manhattan"},"types":[]},
{"neighborhood","political"}]},{"long_name":"Manhattan","short_name":"Manhattan"},"types":[]},
{"political","sublocality"},"sublocality_level_1"},{"long_name":"New York","short_name":"New York"},"types":[]},
{"long_name":New York County","short_name":New York County"},"types":[]},
{"administrative_area_level_2","political"}(...)
"longAddress":163 William Street,
New York, NY 10038, USA"},"nickname":"163 William Street","uuid":"3938134a1b86-4f87-8ad0-f29c66ea674d","longitude":74.00613835924165, "shortAddress":"163 William Street"}
```

This “reverseGeocode” protocol was able to accurately identify the user’s current location as “163 William Street” without this exact address being explicitly provided. The application then used this geocode to find Uber vehicles in the area. We can deduce that a user’s location is saved as well as the vehicles in the area regardless if an Uber is requested for a pick-up. Even if the “reverseGeocode” was not provided, the previous “vehiclepaths” could be mapped out to determine the general location of where the Uber customer stood by seeing where nearby Uber vehicles were pinging.

Privacy Analysis with BlackLight

The above string was found in a file labeled “profiles” under the following folder pathway in the UberClient application folder:

```
Library>Application Support>PersistentStorage>Store>PaymentBaseStreamModelKey
```

The string displays two payment methods that were saved on the Uber app. The first being Apple Pay, this method was not set up on the target device. The second being PayPal, which was the connected method of payment when requesting rides. The redacted portions represent areas where the user’s PayPal login username was saved in plain text. Other fields for “cardExpiration as well as ”cardNumber” are also seen.

As part of our static analysis, we analyzed searches for trips and actual trips using the Uber app on an iPhone. Using BlackLight, we determined that all the locations of pick-ups and drops offs were found in a file called database.db located here:

```
Root>Mobile>Applications>UberClient>database.db
```

In the first experiment, the Uber user searched for an address in Brooklyn (3xx Gold St, Brooklyn, New York 11201). The user decided to take a NYC Yellow Cab instead of using Uber, yet the app tracked the user’s trip with a competitor service, as noted in Figure 7. Interestingly, using Blacklight, we can determine that the Yellow Cab journey was recorded in a very similar way as an Uber journey, as shown in Figure 8.
In the second experiment, the user used the Uber app to find the cost of a journey. The Uber app was then closed. The user then used the Lyft app (competitor service) to see if the journey was cheaper than Uber. Given the cost saving, the user selected the ride with the Lyft app. The user took a Lyft ride from the Cheesecake Factory (100 Cambridge Side Place, Cambridge, MA 02141) to Harvard University. Using Blacklight, we discovered that this Lyft ride was recorded in the Uber app, as shown in Figures 9 and 10. Furthermore, we found that the Uber app was tracking the user for approximately 11 minutes after the Lyft ride had ended, thereby negating Uber’s claim that a user will only be tracked for up to 5 minutes after the conclusion of a ride, although the ride was with a competing service. We proved this by reviewing the route, recorded by the app’s Google API, after the ride had ended and determined that it took the user 11 minutes to complete the mapped route.

5. FUTURE EXPERIMENTS

Our research discovered that the Uber app uses Crashlytics. This third-party analytics service, which was acquired by Google, provides the app developer with crash analytics. Theoretically, while this service is providing real-time crash analytics, it could also be collecting the location of the user device. During our dynamic app analysis with Debookee, we discovered that there were continuous HTTPS requests going back and forth between the device and the Crashlytics server(s), even though the app was not experiencing any crash problems. Additionally, a more extensive examination of comparable “ride” apps would be appropriate to identify whether there are simply privacy concerns with Uber and not with its competitors. There are suggestions that iOS 11 will limit geolocation tracking by third-party app providers. It will be interesting to see how this will impact the Uber app and its tracking abilities.

6. CONCLUSIONS

In an age where many people, especially in the United States, are concerned about government agencies, like the NSA, collecting vast quantities of PII and geolocation information, consumers should also understand the data collection practices of companies, including Uber. Even though Uber placed blame for their app accessing locational services after 5 minutes on Apple’s iOS Maps, it is clear from our findings that their application could store locational data after 5 minutes. With the application able to access location from the iOS Map issue they have cited, the Uber app may still then be able to tap into a user’s location past the 5 minutes indicated in their privacy agreement. The Uber mobile application is saving this location data from vehicles in a user’s area locally to the user’s device. In addition, user’s exact address locations are being determined and locally saved through longitude and latitude coordinates. It is clear that Uber is not just saving trip locations from completed rides, they are collecting geolocation data when the app is not being used for a ride and, more interestingly, is being used to monitor rides with competing services. From a forensics perspective, this research demonstrates the tremendous potential for criminal investigators to use the Uber app to track a suspect and/or victim. The value of this information cannot be overstated given the detailed unencrypted user information available, in plaintext, using forensic imaging tools, like BlackLight.

7. REFERENCES


Appendices and Annexures

Figure 1. Uber dialog box during installation

Figure 2. Screenshot of Google Maps in the Uber app
Figure 3. Google Maps API identified in a PCAP captured by Wireshark

![Google Maps API identified in a PCAP captured by Wireshark](image)

Figure 4. HTTPS requests initiated by Uber

<table>
<thead>
<tr>
<th>Time</th>
<th>URL</th>
<th>Method</th>
</tr>
</thead>
</table>

Figure 5. File in Uber application called "eyeball"

<table>
<thead>
<tr>
<th>Name</th>
<th>Date Created</th>
<th>Date Modified</th>
<th>Date Accessed</th>
<th>Date Added</th>
</tr>
</thead>
<tbody>
<tr>
<td>city</td>
<td>2017-06-14</td>
<td>2017-06-14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>client</td>
<td>2017-06-14</td>
<td>2017-06-14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>clientStatus</td>
<td>2017-06-14</td>
<td>2017-06-14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>eyeball</td>
<td>2017-06-14</td>
<td>2017-06-14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>targetLocationSynced</td>
<td>2017-06-14</td>
<td>2017-06-14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 6. Vehicle path displayed in the Uber application

![Vehicle path displayed in the Uber application](image)
Figure 7. Map in Uber application tracking user in a Yellow Cab

Figure 8. “PICKUP” and “DROP OFF”, with a Yellow Cab, recorded by the Uber app

Figure 9. “PICKUP” and “DROP OFF”, with the Lyft app, recorded by the Uber app
Figure 10. "PICKUP" and "DROPOFF", with the Lyft app, recorded by the Uber app