

Testimony of Ashkan Soltani¹
Independent Privacy Researcher and Consultant

United States Senate, Judiciary Subcommittee on Privacy, Technology and the Law
Hearing on
Protecting Mobile Privacy: Your Smartphones, Tablets, Cell Phones and Your Privacy

May 10, 2011

Chairman Franken, Ranking Member Coburn, and the distinguished members of the Subcommittee, thank you for the opportunity to testify about mobile privacy and the state of location tracking.

My name is Ashkan Soltani. I am a technology researcher and consultant specializing in consumer privacy and security. I have more than 15 years of experience as a technical consultant to Internet companies and federal government agencies. I received my masters degree in Information Science from the University of California at Berkeley, where I conducted extensive research and published two major reports on the methods by which users are tracked online and to what extent. Last year, I served as a staff technologist in the Division of Privacy and Identity Protection at the Federal Trade Commission on investigations related to Internet technology and consumer privacy. I have also worked as the primary technical consultant on The Wall Street Journal's *What They Know* series investigating issues relating to privacy online.

Recent revelations about how mobile devices handle sensitive data—particularly location information—have surprised consumers. Their devices often play a large role in their everyday activities, and many consumers show significant concern about who has access to their information.² Whether consumers understand these privacy risks and whether they have meaningful control over information access are critical questions for this Subcommittee.

I have been invited to testify about the current state of mobile privacy and location tracking from a technical perspective. First, I will describe location-based services and how a mobile device can determine its location. Second, I will discuss three recent issues that demonstrate how location data and other personal information are collected and shared in the current mobile ecosystem. Finally, I will discuss three broad implications for consumer mobile privacy and provide some suggestions for improvement.

¹ My oral and written testimony here today to the Subcommittee represents my own personal views, and does not reflect the views of any of the organizations that I have consulted or worked for in the past.

² Tsai, Janice Y., Kelley, Patrick Gage, Cranor, Lorrie Faith, Sadeh, Norman. Location Sharing Technologies: Privacy Risks and Controls. (2009). From <http://repository.cmu.edu/isr/85/>

A. MOBILE DEVICES AND LOCATION-BASED SERVICES

Mobile devices today are powerful computing machines. Like desktop computers, many mobile devices run complex operating system platforms that allow third-party developers to create software applications to perform specialized tasks. Two of the most widely used mobile platforms, Apple iOS and Google Android, offer consumers hundreds of thousands of innovative applications to download and install onto their devices through the Apple App Store and Android Market. These include e-mail capabilities and productivity tools, mapping and navigation services, social media applications and games. However, unlike desktop computers, mobile devices are uniquely *mobile* which introduces unique privacy implications for their owners.

Consumers take their mobile phones and tablet computers with them nearly everywhere they go.³ They often carry these devices in their pockets from their homes to their offices, while traveling by car or train, when on their way to daycare and to the grocery store. Mobile phones, in particular, are personal “always-on” devices; therefore, the location of these devices often closely mirrors that of their owners’ locations and activities.

The location of a mobile device at any given moment may not be particularly sensitive; However, the historical trail of past locations can reveal much about its user’s behavior. In some cases, a person who has access to historical location data can infer trends that uniquely identify an individual. For example, if a mobile device’s location is the same each work day, then consistently at another location every evening, it might expose the location of the device owner’s workplace and home, respectively. An individual or organization with access to this information could then correlate it with public databases that could then be linked to a particular individual.⁴

However, location-based services (LBS) are a major selling point for many mobile devices. These features quickly enable the discovery of nearby stores and restaurants, sharing of current location with friends and family by using “check-in” functionality within social networking applications, and easy directional navigation to desired destinations. In order to provide this functionality, the application or service provider needs to pinpoint and use the mobile device’s location.

³ Three in five mobile phone owners say they carry their phones at all times, even inside the home. See: Stanton, D. (2008, September 8). New Study Shows Mobile Phones Merging New, Established Roles. Knowledge Networks. From http://www.knowledgenetworks.com/news/releases/2008/091808_mobilephones.html

⁴ Golle, Philippe and Kurt Partridge. On the Anonymity of Home/Work Location Pairs. From <http://xenon.stanford.edu/~pgolle/papers/commute.pdf> (Researchers demonstrate it may be possible to associate home/work location pairs to individuals’ identity.)

Note: The icons in the margins below refer to the diagram in Appendix A and are used to direct attention to specific portions of the “Location Ecosystem.”

There are four primary ways the location of a mobile device can be determined, depending on both its hardware and software capabilities.



1. Global positioning system (GPS) is a technology that allows a device to determine its location by triangulating GPS satellite signals, which are typically accurate to within a few meters. While nearly all smartphones manufactured today contain a built-in GPS chip, many mobile devices (*e.g.*, laptops) typically do not. While GPS allows for high accuracy of location, it is often unavailable indoors and its high consumption of battery life often compels users to turn off GPS until they require it.



2. Wireless carriers can help mobile devices determine location by using information about the signals of nearby cell phone towers. This is called cellular geolocation. Cellular phone towers act as known “landmarks” since they have fixed locations. This property enables wireless carriers to triangulate a device’s location anytime the device is powered on. Mobile phones can send a query to the carrier to request the physical coordinates of towers within range and then calculate its position as best as possible. This technique is generally less accurate than GPS and varies widely depending on the density of cell towers in a given area.



3. Location providers are services that allow devices to determine location via a variety of methods, which include cellular, Wi-Fi and Internet Protocol (IP) based methods. Companies such as Google, Apple, and Skyhook can act as location providers by compiling extensive databases that correlate Wi-Fi access points and cell phone towers with their physical locations. Mobile devices then query these databases with information about nearby “wireless landmarks” (*i.e.*, Wi-Fi access points and cell phone towers) in order to obtain their current location. As a result, the location provider is able to infer the current location of the mobile device as well as enhance its own location database with any additional “wireless landmarks” provided with the query.



4. Location aggregators are a separate class of location service providers that obtain location information via direct arrangements with wireless carriers. As such, device location is obtained directly from triangulation of nearby cellular tower data and does not rely on the handset to be “aware” of its present location. This enables features such as ‘geofencing,’ which is the ability to notify a third party whenever a device enters geographic area without requiring a specialized application on the phone. Location aggregators occupy a unique niche in this marketplace as they have a detailed “carrier view” vantage point across all of their participating partners, and they provide data to third party applications and websites directly.

B. HOW MOBILE DEVICE LOCATION IS COLLECTED AND SHARED



1. By Location Providers

The process by which Location Providers gather data raises significant privacy concerns. Much of the initial public concern focused on Google's reported collection of consumer information when it mapped wireless landmarks like cell towers and Wi-Fi access points by using employee-driven automobiles that were equipped with special sensors.⁵

More recently, location providers began distributing the work by using their customers' mobile devices as "scouts in the field" in order to compile their databases of the physical locations of wireless landmarks. This "crowdsourcing" of location data has introduced additional privacy concerns. By leveraging consumers' mobile devices as scouts, location providers consequently receive the location of the mobile device as they report their findings.⁶ Consumers have the option to "opt-out" of this practice; however, background collection and transmission of location information is enabled by default for most location providers.⁷

Even the notice that is offered may also be inadequate for meaningful choice. Figure 1 below compares the Google Android platform's permission screen informing users of the background collection of location data to the comparable screen on the Apple iOS platform. A customer would have to read Apple's lengthy software license agreement to learn that disabling location services means disabling the background collection of location data.

In addition, a mobile device user's attempt to "opt-out" may be ineffective. In April 2011, The Wall Street Journal reported that Apple iPhone devices would still collect and transmitting this information, even when users' had affirmatively set the location services to "off." That is, even when consumers elected to disable collection of their device location, their iPhones had continued to record and transmit location services information to Apple's servers.⁸ Surprisingly, this scenario conflicts with a July 12, 2010 letter from Apple's General Counsel to Representatives Ed Markey and Joe Barton which stated that "Apple automatically collects this

⁵ Stone, Brad. (2010, May 14). Google Says It Collected Private Data By Mistake. From <http://www.nytimes.com/2010/05/15/business/15google.html>

⁶ Valentino-Devries, Jennifer. (2011, April 23). Google Defends Way It Gets Phone Data. From <http://online.wsj.com/article/SB10001424052748703387904576279451001593760.html>

⁷ Google's default is enabled by means of a pre-selected check box during the initial product setup which a user has to actively 'uncheck'. See Figure 1. The FTC has raised concerns about "pre-checked" dialogues as a mechanism for affirmative consent in a recent settlement with Google and their Buzz social networking product. See <http://www.ftc.gov/os/caselist/1023136/110330googlebuzzcmpt.pdf> at page 4.

⁸ Valentino-Devries, Jennifer. (2011, April 25). iPhone Stored Location in Test Even if Disabled. From <http://online.wsj.com/article/SB10001424052748704123204576283580249161342.html> and Apple. (2011, April 27). Apple Q&A On Location Data. From http://www.apple.com/pr/library/2011/04/27/location_qa.html.

information only (1) if the device's location-based services capabilities are toggled to 'On' and (2) the customer uses an application requiring location-based information.”⁹

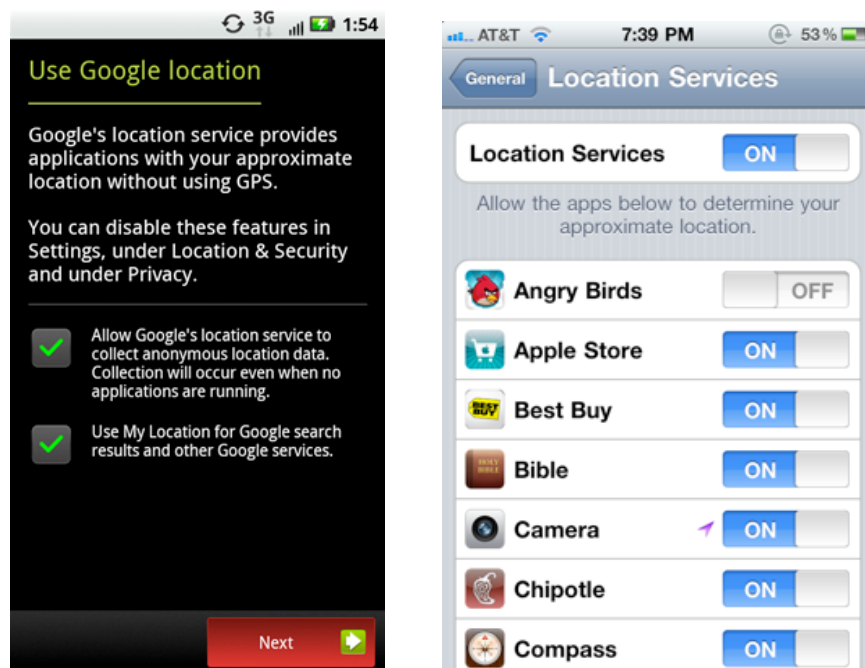


Figure 1. Permission screens controlling location service on the Android and iPhone platforms. (Location Services and subsequent collection is ON by default on both platforms.)



2. In Local Cache Files on the Device

In order to improve the speed of location look-ups and to further reduce battery consumption, many mobile platform developers design their systems to keep a local copy - a “cache” - of location information from previous queries on the mobile device. This allows a mobile device to determine its location without having to re-query the location provider every time it’s near a previously seen landmark.

Like any repository of sensitive information, this cache of location data poses potential privacy issues. As mentioned previously, a person who is able to gain access to this database might be able to determine the user’s past whereabouts (subject to the historical length of the cache). In addition, last month, researchers identified a cache of location data that includes a full year’s worth of location history stored on their Apple iPhone device.¹⁰ This data had been recorded by

⁹ Apple Inc’s Response For Information Regarding Its Privacy Policy and Location-Based Services. (2010, July 12). From <http://markey.house.gov/docs/applemarkeybarton7-12-10.pdf> at page 7.

¹⁰ Allan, Alasdair and Pete Warden. (2011, April 20). Got An iPhone or 3G iPad? Apple Is Recording Your Moves. From <http://radar.oreilly.com/2011/04/apple-location-tracking.html>.

iPhones even when a user elected to disable location services. This effectively means that, in addition to there being no meaningful mechanism by which consumers can disable the background collection of location data by location providers, they also lack a meaningful mechanism to disable the collection of location data in a cache file. The researchers also found a copy of this same cache file stored insecurely on computers that had been used to synchronize or backup their iPhones, iPads, and other iOS devices.¹¹

By analyzing the data stored in this cache, which is a record of nearby cellular towers and Wi-Fi access points the phone encountered, the researchers were able to re-create a map of their previous travels from Washington DC to New York, as shown below in Figure 2. They also publicly released a tool that consumers could use to easily access and visualize their own location histories.¹²

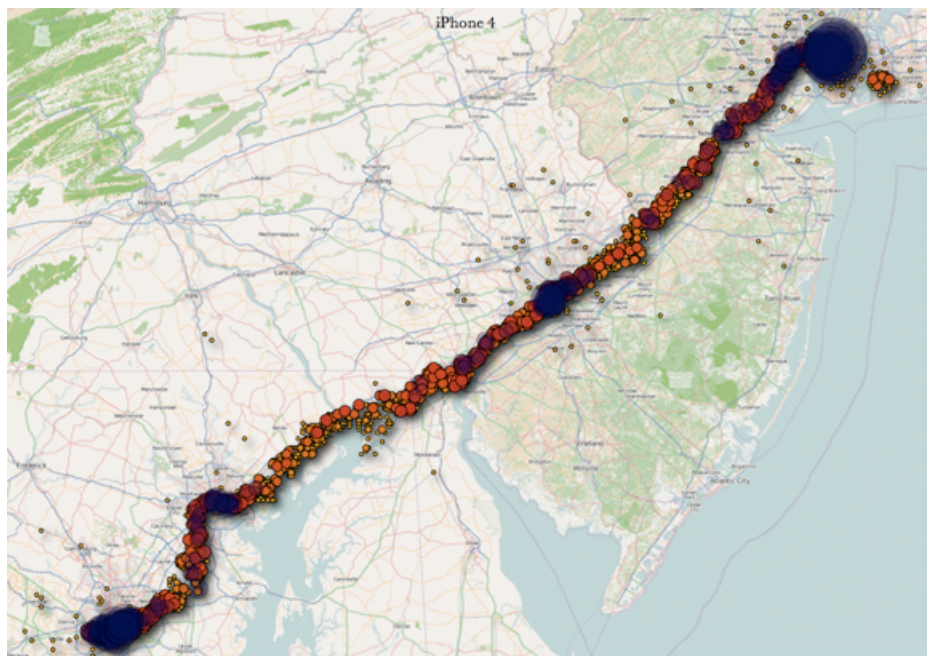


Figure 2. Map of researchers' whereabouts, inferred from local iPhone cache.¹³

¹¹ Apple announced a fix for this bug which reduces the size of the location database cache, stops transfer to iTunes when you connect your device to a computer, and deletes the cache entirely when you turn Location Services off. However, this fix doesn't apply to older 2G and 3G devices. Chen, Jacqui. (2011, May 05). iOS 4.3.3 is out with location tracking fixes for iPhone, iPad. From <http://arstechnica.com/apple/news/2011/05/ios-433-is-out-with-location-tracking-fixes-for-iphone-ipad.ars>

¹² Warden, Pete. (2011, April 20). iPhoneTracker. From <http://petewarden.github.com/iPhoneTracker/>

¹³ Allan, Alasdair. (2011, April 20). Got an iPhone or 3G iPad? Apple Is Recording Your Moves. From <http://radar.oreilly.com/2011/04/apple-location-tracking.html>

Further research into competing platforms showed that Apple was not alone in this practice. Google and Microsoft smartphones also cache location histories, although the retention period for this information on these platforms appears to be shorter.¹⁴

It's worth noting here that while the recent "discovery" of local location caches has better informed the public about the issue, researchers and law enforcement have been aware of this practice for some time.¹⁵ In addition to location history, researchers have repeatedly demonstrated that personal information such as email, text messages, browsing history, photos, and passwords can be recovered easily with physical access to the devices and, in some cases, remotely.¹⁶ Surprisingly, this is even true for applications typically thought to be impervious to monitoring, such as the encrypted voice calling program Skype.¹⁷

3. By Smartphone Applications



In addition to storing location data locally and transmitting it to Location Providers, many users' smartphones will transmit their location and other sensitive data to numerous third parties via the use of third-party applications, such as games and other software programs. The specific parties and amount of information will vary depending on the specific "apps" used. However, the practice of transmitting potentially sensitive data off of the device is common for most applications.

¹⁴ Gohring, Nancy. (2011, April 29). Microsoft Admits To More Windows Phone Update Problems. From http://www.pcworld.com/article/226733/microsoft_admits_to_more_windows_phone_update_problems.html and Foresman, Chris. Android Phone Keeps Location Cache Too, But It's Harder To Access. From <http://arstechnica.com/gadgets/news/2011/04/android-phones-keep-location-cache-too-but-its-harder-to-access.ars>

¹⁵ Levinson presented his research on the iPhone cache file at a conference six months ago and subsequently published his findings in December 2010. Levinson, Alex. (2011, April 21). Three Major Issues with the Latest iPhone Tracking "Discovery." From <https://alexlevinson.wordpress.com/2011/04/21/3-major-issues-with-the-latest-iphone-tracking-discovery/>. Johnson, Bobbie. (2011, April 21). Researcher: iPhone Location Data Already Used By Cops. From <http://gigaom.com/2011/04/21/researcher-iphone-location-data-already-used-by-cops/>.

¹⁶ Edwards, Sarah. Inside the App: All Your Data are Belong to Me. From <http://www.shmoocon.org/speakers#insideapp>

¹⁷ A design vulnerability in the secure calling software Skype allows access to "full name, date of birth, city/state/country, home phone, office phone, cell phone and email addresses" of users because files on the device had insecure permissions and we stored in an unencrypted format. Case, Justin. (2011, April 15). (Updated) Exclusive Vulnerability In Skype For Android Is Exposing Your Name, Phone Number, Chat Logs, And A Lot More. From <http://www.androidpolice.com/2011/04/14/exclusive-vulnerability-in-skype-for-android-is-exposing-your-name-phone-number-chat-logs-and-a-lot-more/>

In a survey of the 101 popular iPhone and Android phone apps in December 2010, The Wall Street Journal found that 47 of them transmitted the phone's location and 56 also transmitted identifiers (such as hardware serial numbers) to third parties.¹⁸ Sometimes this information would go to the application developer's server, such as Yelp.com when using the Yelp "app." Other times, the location would be shared by the app further afield to its advertising partners without clear indication to the end-user. Forty-five apps had no discernible privacy policies, and neither Apple nor Google requires apps to have privacy policies.

While user consent is typically required before applications are allowed to access location information, the purpose may not always be apparent to the user, and the user may have no indication that this information will subsequently be disclosed to third parties. For example, one iPhone app called Ninjump—a game—accesses and sends the a mobile device's location information to its mobile ad provider.¹⁹ Most users would probably be befuddled about why an action game would ever need to access their location or disclose it to others, even if they consented to the initial collection of this information.

Data sharing isn't limited to location information. Applications can access and transmit data which includes text messages, emails, phone numbers, contacts stored, and even browser history stored on the device, as well as any information users knowingly enter in the process of using the app.²⁰ Some of this sharing may be expected, while other times it may be surprising. One example is where a popular social networking application had uploaded entire copies of users' address books to Facebook's servers.²¹

C. IMPLICATIONS FOR CONSUMER PRIVACY

These recent issues demonstrate key points of contention between consumers privacy and business interests.

1. Existing Notice and Choice Mechanisms Are Insufficient

Mobile apps and platforms do not provide consumers with sufficiently detailed notices about how their location and other sensitive information will be collected and used. Notice requirements vary from platform-to-platform. However, many disclosures related to privacy, such as data retention and sharing, frequently go unmentioned. The notices also rarely differentiate between first and third party data uses nor do they reveal business partners, like ad

¹⁸ Thurm, Scott and Yutari Iwatani Kane. (2010, December 17). Your Apps Are Watching You. From <http://online.wsj.com/article/SB10001424052748704694004576020083703574602.html>

¹⁹ WSJ Blogs. (2011, December 17). What They Know Mobile. Ninjump. From <http://blogs.wsj.com/wtk-mobile/2010/12/17/ninjump/>

²⁰ Seriot, Nicolas. (2010). iPhone Privacy. From http://seriot.ch/resources/talks_papers/iPhonePrivacy.pdf

²¹ Moos, Kurt von. (2010, February 26). Privacy Fails: How Facebook Steals Your Friends Numbers. From http://kurtvonmoos.com/facebook_steals_contact_info/

networks, by name. As such, consumers are unable to make meaningful choices regarding their privacy risks when using mobile devices.

For example, with the exception of real-time location data, Apple's iOS platform for iPhones does not disclose to users what other location information may be accessed and shared by applications upon download. The iOS platform also does not inform users if an app will collect information from their address books, calendars, or other data from their iPhone.

Consumers are given a chance to "click through" to discover individual app privacy policies, but these are often long legal statements that are particularly difficult to read on a small mobile screen,²² when they're even available. Comparatively, the Android platform allows more descriptive notices informing users of the data an app will collect. Although many of the terms used in these notice are still very technical in nature and can appear cryptic for a lay user to understand.

While mobile platforms today allow users to first review these disclosure notices before they install an app. But they also all adopt a "take it or leave it" approach to application permissions: the user can either allow access to all of the information the app requests, or deny all access (and thus not install the app). Granular permissions are not typically made available. That is, users are forced to give up their location information if they want to play the Ninjump game.

2. Collected Location Information Can Be Sensitive

Some industry players dismiss the recent concern about location privacy by saying that the information collected is not actually *device* location information. In Apple's Q&A on location data, they say that some of the collected information is about network equipment "some of which may be located more than one hundred miles away."²³

While this may be true for cellular location in sparse rural areas, many urban environments yield device location measurements as accurate as 50 to 200 feet.²⁴ Since Wi-Fi is a short-range communication, knowing even one nearby Wi-Fi signal can typically pin the user within 100 feet.

²² This matter became the underlying premise of a popular television show parodying "Apple's ridiculous 55-page iTunes terms and conditions." O'Grady, Jason. D (2011, April 28) South Park parodies iTunes terms and conditions. From <http://www.zdnet.com/blog/apple/south-park-parodies-itunes-terms-and-conditions/10043>.

²³ Apple. (2011, April 27). Apple Q&A On Location Data. From http://www.apple.com/pr/library/2011/04/27/location_qa.html

²⁴ Steve Lee, product manager for Google Maps for Mobile and Google Latitude said in a May 2010 email that Google had 300 million Wi-Fi networks in its database which could pinpoint a device's location to within about 100 feet. Efrati, Amir. (2011, May 1). Google Calls Location Data 'Valuable.' From <http://online.wsj.com/article/SB10001424052748703703304576297450030517830.html>

As a quick demonstration, I recorded my device's location while sitting on a bench in the lobby of Hart Senate Office Building. Using GPS, my location was accurately reported to within 20 meters, as indicated by the small circle at the center of the left image in Figure 3 below. The right image shows nearly the exact same location found using Wi-Fi geolocation, which only uses a location database maintained by Google.

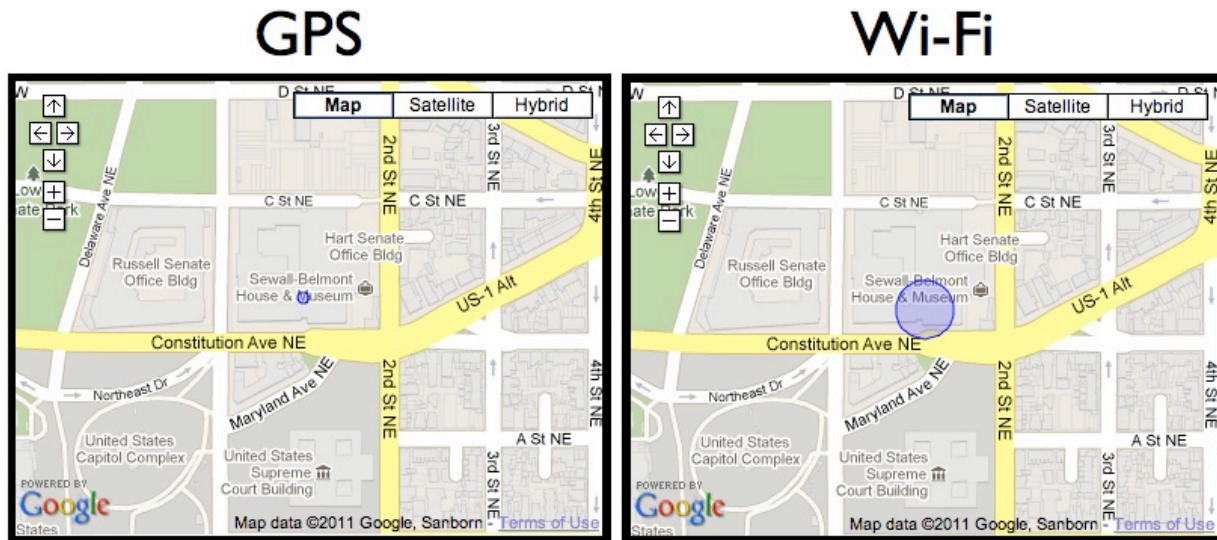


Figure 3. Comparing the accuracy of GPS and Wi-Fi based geolocation techniques.²⁵

Quite a lot of information can be deduced from trails of historical location data. People are creatures of habit,²⁶ and it would often be easy to deduce where an individual works from her location on weekdays from 9am-5pm or, from the same nightly location, where she sleeps. These two pieces of information start to form a picture of who the device owner is.

3. Location Data Can Be Tied to Consumer Identities

Industry also argues that location data cannot be associated with consumers' real identities, and that this data is often simply "anonymous usage statistics."²⁷ However, to the degree that this data is also associated with unique identifiers—such as serial numbers or IP addresses that can

²⁵ The strongest Wi-Fi signal my device could detect was one of the "Odyssey" access points. Google's geolocation database reported the location of this access point (and thus my location) to within 36.0 meters as indicated by the circle in the right image (as identified by Google).

²⁶ 93% of people return to the same locations: Song, C., Qu, Z., Blumm, N., and Barabási, A.L. Limits of predictability in human mobility. *Science*. 2010 Feb 19, 327(5968): 1018-21. From <http://www.ncbi.nlm.nih.gov/pubmed/20167789>

²⁷ "Google spokesman said it collects information anonymously." Kane, Yukari Iwatani, and Jennifer Valentino-DeVries. (2011, April 28). Jobs Tries to Calm iPhone Imbroglia. From <http://online.wsj.com/article/SB10001424052748703367004576288790268529716.html>

later be linked back to an individual device or person²⁸—it becomes difficult to refer to it as “anonymous information.”²⁹

Identifiers enable further correlations with additional information generated via other channels, such as subscriber information (from a wireless carrier), login credentials (from phones that sync their e-mail or calendars), or even in some cases name, credit card or address information used in the app marketplace. For example, research recently demonstrated that “anonymous” device identifiers can easily be correlated to user’s location and identity” in the form of pseudonyms and Facebook profiles with a reasonable degree of likelihood.³⁰

Whether re-identification is possible depends on what other information is available, which itself hinges on the data retention and security practices of multiple participants in this ecosystem. It is rarely the case that information should be called “anonymous,” since there is nearly always some small chance of re-identification.

Fortunately, at least some in industry share this view. When asked about the anonymity of location, the CEO of Location Provider Skyhook Jay Yaroo stated:

“I[f] you associate any history of a user at all it’s very easy to, after the fact, figure out the name of that user. So when you hear companies like Microsoft and Google say, ‘We’re anonymizing the data,’ it doesn’t matter. If there’s a location history, all I do is look at past 9 o’clock and there’s a 95% chance that you went home. And I will look at that, and I will look up that address and I will know who you are. And as you start adding more and more data, I match that with where you work and now I know this is you.”³¹

²⁸ While IP addresses can be dynamic, they can persist for days. IP addresses assigned to phones on the Verizon and Sprint do not change over a 2-day test period. See Balakrishnan, Mahesh, Iqbal Mohamed, and Venugopalan Ramusubramanian. (2009). Where's That Phone? Geolocating IP Addresses on 3G Networks. From <http://research.microsoft.com/en-us/um/people/maheshba/papers/ephemera-imc09.pdf>

²⁹The Dutch Data Protection Authority argues that MAC addresses, in combination with the ability to identify the location of wireless hardware, may by itself qualify as personal information. Preuschat, Archibald. (2011, April 20). Google Faces New Demands In Netherlands Over Street View Data. From

<http://online.wsj.com/article/0,,SB10001424052748703922504576273151673266520,00.html>

³⁰Recently, a researcher demonstrated that device IDs can be linked to GPS location (30%), Weak Identities (20%), and Facebook profiles (10%) using public game service OpenFeint. See Cortesi, Aldo. (2011, May 4). De-Anonymizing Apple UDIDs with OpenFeint. From <http://corte.si/posts/security/openfeint-udid-deanonymization/index.html>

³¹ Yaroo, Jay. (2011, April 28). Everything You Need To Know About How Phones Are Stalking You Everywhere. From <http://www.businessinsider.com/skyhook-ceo-2011-4>

D. Conclusion

As mobile devices become more powerful—and more ingrained in the way consumers work and play—information about where a device is located becomes an ever more valuable input for commercial activity. But at the same time, consumers have expressed significant concern about how their devices expose sensitive information about them in ways they might not expect. Consumers need to be able to trust their devices in order to take full advantage of all the benefits mobile technology has to offer.

To better protect consumer privacy going forward, I offer four suggestions:

1. Mobile platform providers and application developers should work together to provide consumers with more transparency into exactly what data are collected, how they are stored, to whom they are transmitted, and how they are secured and used.
2. Certain disclosures should be mandatory, such as clearly differentiating between first and third party uses of all potentially sensitive data, and also between active use and passive background activity. Precise definitions for "location" and "identity" should be provided.
3. Providers and developers should also work to ensure that the information consumers entrust with them are handled securely and in line with their expectations.
4. Providers and developers should also offer meaningful choice, such as granular permissions and working opt-outs, to consumers so they can make effective, privacy-conscious decisions in the marketplace.

Thank you for the opportunity to testify here today. Mobile privacy is a very nuanced issue, even for us technologists, so I thank the subcommittee for their attention on this increasingly important problem. I will be happy to answer any further questions.

Appendix A: Flow of Location Data in Mobile Ecosystem

