Location is Everything

Balancing Innovation, Convenience, and Privacy in Location-based Technologies

SUBMITTED BY:

C. Christopher Post
Stephen Woodrow
{ccpost,woodrow}@mit.edu

FOR 6.805/STS.487:
ETHICS AND LAW ON THE ELECTRONIC FRONTIER

Submitted 10 December 2008
(Revision 83)

Abstract

Location-based technologies (LBTs) — services, devices, and appplications that provide functionality and content tailored by knowledge of your current location — are becoming popular with consumers and are expected to be a major growth area in the mobile device market. While offering great convenience, LBTs are implemented significantly differently from previous technologies providing similar features, and these differences pose new privacy risks to users of LBTs. This paper shows that users of new LBTs expose themselves to privacy risks that are difficult to understand and manage, due to the nature and current implementation of most LBTs.

We begin by identifying and discussing the essential differences between LBTs and previous technology efforts to gain an understanding of how LBTs differ from previous technologies, and the privacy implications of these differences. We proceed to explore how these issues manifest themselves in implemented LBT systems commonly available today, including the Apple iPhone, the T-Mobile G1, and the Skyhook Wireless geolocation system. Finally, with an understanding of some of the theoretical and real pitfalls in LBTs, we propose measures to improve user privacy. These measures include new law to provide accountability for data collection and use, policy approaches to ensure users understand their privacy risks and have real choice in managing these risks, and technical methods to deal with some particulary concerning problems we uncovered.

LBTs are not a novel fad — they do provide valuable functionality, especially for users of mobile devices. As such, we cannot simply expect users to reject LBTs because they cannot determine if a particular LBT is compatibile with their personal privacy values. In order to promote continued innovation in this market, improvements in law, policy, and technology are required to allow users to enjoy the convenience of LBTs while also maintaining a real expectation of privacy when using these technologies.

Acknowledgements

This paper was a long and complicated effort, and the authors would like to thank the following people for their help:

- Jane Horvath, Senior Privacy Counsel, Google Inc., for speaking with us about Google's approach to privacy and specific aspects of Google's location-based services.
- Hal Abelson, 6.805 Professor, for forcing us to think long and hard about the thesis of our paper, offering feedback on our draft, and for general wisdom and contacts.
- Shekhar Krishnan, 6.805 TA, for also forcing us to think long and hard about the thesis of our paper, and for offering feedback on our draft and other submissions.
- Mike Fischer and Les Perelman, readers and reviewers of our early ideas.
- Kipp Jones of Skyhook Wireless, as well as Hari Balakrishnan and the other readers of `csail-related` who responded to our request for contacts at Skyhook.

About the title

The phrase "Location is everything" is a maxim of real estate, presumably related to the maxim coined by British real estate developer Harold Samuel: "There are three things you need in property, these are: location, location, and location."[1]

[1] `http://www.fundinguniverse.com/company-histories/Land-Securities-PLC-Company-History.html`

Contents

List of Figures

Introduction

Location-based technologies (LBT) — services, devices, and applications that provide functionality and content tailored by knowledge of your current location — are about to become the "next big thing" for mobile device users. While some of these systems have been around for several years, they are just now beginning to gain traction with the widespread availability and popularity of devices, like the Apple iPhone, that support and promote these services and applications. And this is just a start — one analyst predicts 3000% growth in this area, leading to a $15B/year business over the next 5 years [3]. As the growth of the market for these devices continues to explode, new legal, policy, and technical controls are required to allow users of location-based services to enjoy the convenience and benefits of the service while feeling informed and safe in their privacy choices.

While mobile device users may have some experience with basic location-based devices like GPS receivers or in-car navigation systems, these new LBTs, and in particular location-based services and applications, are fundamentally different from these services despite similar functionality or appearance. Unlike GPS, a passive geolocation system, LBTs involve two-way communication between the user and the service's host, and this data is often transferred to several parties, and retained by the parties, to conduct a single transaction. Depending on the service and its business practices, this may expose users to privacy risks, but this shouldn't come as a surprise — there is some risk whenever data is transmitted/shared. However, unlike other privacy risks facing users today, the nature of modern LBTs makes it extremely difficult for users to manage their privacy risks. While some users may choose to avoid these services as a result, it is likely that most will make ill-informed choices about their privacy to enjoy the benefits of LBTs.

Defining Location-based Technology

Location-based technologies (LBT), sometimes also referred to as location-based services (LBS), connote an extremely broad and buzzword-filled field, and can be taken to mean any number of types of services, devices, and applications. In the context of this paper, "location-based technology" refers to applications and services that utilize location information about a mobile user to provide a customized service or content to the user, as well as the mobile devices that host these applications and services. We will also include mobile geolocation services, services that determine the geographic location of a user, typically for the purpose of a LBT application or service.

Paper Roadmap

This paper begins with a few hypothetical vignettes to illustrate some potential privacy concerns with LBT, followed by an overview of commonly-available location-finding and location-based technologies. Then, background is presented on the next generation of location-based technologies to illustrate the differences between the old and new approaches to LBT. From there, an analysis of the privacy risks and privacy mangagement implications of these new technologies will be conducted, using specific technologies in the market to illustrate our arguments. Following this, we discuss and recommend specific approaches to improving user privacy management.

Chapter 1

Scenarios

Given the current technical operation of modern location-based technologies and the vague legal and policy frameworks, there is vast potential for abuse of location information. To date, there have not been any major reported abuses of location information through these new technologies. The services are still in their infancy, however, and could easily suffer substantial problems in the future. It is clear that the architects of many of these services have indeed made privacy an important concern, but there is still a long road ahead.

In this section, we present hypothetical scenarios that are plausible given the current legal, policy, and technical architecture surrounding location-based technologies. While they may sound extreme, they intentionally make use of many of the flaws considered throughout Chapter IV. Scenarios from two categories will be considered: the government and third parties.

1.1 The Government

The U.S. Drug Enforcement Administration (DEA) has identified Port City, U.S.A. as a major distribution point for illegal drugs. Its coastal location means that the fishing business thrives there, and there are numerous private and commercial marinas near the city with hundreds of boats entering and leaving port every day. The bustling marine business has also attracted business in the illegal drug import trade. The DEA has attempted to find drugs as they enter the country on watercraft, but the vast number of boats makes this method highly ineffective.

Upon recent investigation, however, the DEA has identified two specific houses in the rural areas surrounding Port City that it strongly believes are distribution points for the incoming drugs. With the recent tough economic times, the DEA has been unable to devote sufficient financial resources to physically conduct surveillance on the houses to find further distribution points. Thus, the DEA decides to take a more technical approach to the problem. It is able to issue a court order to a major location service provider based on reasonable suspicion from very limited physical surveillance to retain the records (including unique identifier of each user and time/date) of any location queries that result in a location fix within a small radius of both of these houses.

Within a few weeks, the DEA has identified a clear patter in one user's location fixes. Every few days, he visits one or both houses, usually in the evenings. He also regularly visits a third house. Naturally, the DEA concludes that this user must be part of the drug trafficking ring, and that there is also a third house involved in the distribution! Based on this information, the DEA decides that there is a "substantial chance" or "fair probability" per *Illinois v. Gates* (1983) that this user and the third house are participating in the drug ring [4]. This gives the DEA probable cause to obtain a warrant to search both the third house and the apparent residence (as shown in location queries) of the user who visited all three houses.

Upon searching the two houses identified through location data, the DEA finds no evidence of drugs. When questioned, it becomes apparent that the user identified through location data works as a pizza deliveryman, and the only apparent connection between him and the third house is that the third house happens to order pizza very frequently. In fact, the third house is a fraternity house, and gives the pizza business a very steady stream of income. What about the connection between the pizza deliveryman and the two drug houses, though? As it turns out, drug traffickers must also really like pizza.

In this scenario, the government used the weak "reasonable suspicion" standard to obtain location information from location-based technology users who happened to be around the drug houses. This data was then used to infer information about innocent users in order to qualify for the slightly more stringent "probable cause" standard. The vague legal standards regarding location data allowed the DEA to leapfrog to a higher standard for obtaining a warrant, which in turn allowed them to violate the rights of the pizza deliveryman and the fraternity house. Inferences based on location data can be powerful, but they can also be wildly inaccurate in some cases.

1.2 Third Parties

Alice is a single mother of two children who lives in Cambridge, MA. She wasn't always single, though. She also didn't always live in Cambridge, either. A year earlier, she had a husband, Bob, and lived in San Jose, CA. Bob worked for a technology startup in Silicon Valley when he was married to Alice, and spent long hours working as a programmer. While Bob's company was doing well, his marriage was doing well.

One day, though, Bob's company was up for another round of funding from the venture capital firm. Unfortunately, they had not been able to pull in enough revenue from their service to make it apparent that the business was profitable, and the venture capital firm did not renew them for another round of funding. Without funding or a source of income, the company went under and Bob was out of a job. He had spent countless hours trying to make their product a success, but in the end it was not enough. Distraught, Bob turned to alcohol during his unemployment. In order to support the children, Alice had to get a part-time job.

After time, Bob's alcoholic tendencies increased, and it became apparent he was not on the track to finding a new job. He even started physically abusing Alice and their children. At that point, Alice decided that she need to make a drastic change for her and her children: she left Bob and moved across the country to Cambridge, hoping that Bob would never find them again.

Bob, however, was a tech savvy consumer. When Alice lived with him in San Jose, they had a wireless access point, as many modern households do. Alice took this access point with her when she moved to Cambridge. Because the WiFi databases maintained by various location-service providers must keep their databases up to date in order to maintain the accuracy of the service, Alice's access point showed up in the databases as being located in Cambridge. Bob had record of the MAC address of the access point (the BSSID), and was able to query the database of a particular location-service provider to find the location of Alice's access point.

It took a while for the database to update, but being a programmer, it was a simple task for Bob to write a program to automatically query the database and let him know when a location was found. Now, Bob not only knew what city Alice and his children were in, but the exact house in which they now lived. Still angry and an alcoholic, Bob went to Cambridge and began harassing Alice and the two children. Now Alice must either undertake the lengthy process of taking legal action against Bob or move yet again, just because her access point betrayed her location.

Here, Alice's safety and the safety of her children was betrayed by the massive database of WiFi access points kept by location-service providers. In order for the location services to operate, it is indeed necessary for the providers to map the location of Alice's access point. A flaw in the technical operation of this WiFi database allowed a malicious third party to obtain an innocent user's home location.

Chapter 2

History

2.1 GPS

Since the Global Positioning System (GPS) was declared fully operational by the United States Department of Defense in 1995, it has seen widespread use by both military and civilian applications [5]. This was the first system that allowed anyone with a GPS-capable device to near-instantly determine his location on the earth without special training or knowledge. In the most basic sense, a GPS device answers the question “Where am I?” using absolute coordinates (latitude, longitude, and usually elevation). Even before 1995, the partially completed system received widespread publicity thanks to its application in the Persian Gulf War during 1991. The system allowed the allied military to effectively plan and track the movement of forces across a desert landscape that was lacking landmarks for other means of navigation. Being able to determine ones location anywhere on earth proved to be extremely beneficial to the military, and it later showed promising potential for civilian and commercial applications as well.

2.1.1 Technical Operation

GPS functions using a constellation of orbital satellites that constantly broadcast signals that are received by GPS devices. These devices then use the received data to calculate position. Originally, the orbital constellation contained 24 satellites in 6 different circular orbits [6]. The current system, however, uses 31 active satellites in a non-uniform arrangement in order to improve accuracy and satellite failure resiliency. In any arrangement, however, the constellation is designed such that a GPS device can ideally see at least 6 satellites at any time from any point on the earths surface. In practice, however, a GPS device may not be able to reliably receive signals from all satellites it can theoretically see due to signal degradation from buildings or geographic features. As long the device can reliably receive signals from at least 4 satellites, though, it can calculate its position.

In order to compute position, a GPS device listens for and locks on to at least 4 satellite signals. For civilian use, these signals are broadcast on the L1 (1575.42 MHz) frequency, and use a special modulation method to ensure that signals from various satellites do not interfere with each other even though they are on the same frequency [1]. Each satellite broadcasts its specific absolute position as well as a carefully calibrated time signal. All satellites in the constellation continually calibrate their clocks among each other so that they all are running on precisely the same time. Depending on the relative distances of the satellites, the signals are received by the GPS device at slightly different times. Using 3 signals, the device can determine the distance of each of the 3 satellites and compute its position relative to the satellites as shown in Figure 2.1. It then uses the satellites broadcasted absolute positions to determine its own absolute position (latitude, longitude, and altitude). A 4$^{\text{th}}$ satellite signal is used to calibrate the devices internal clock so that it can accurately determine time delays of received signals.

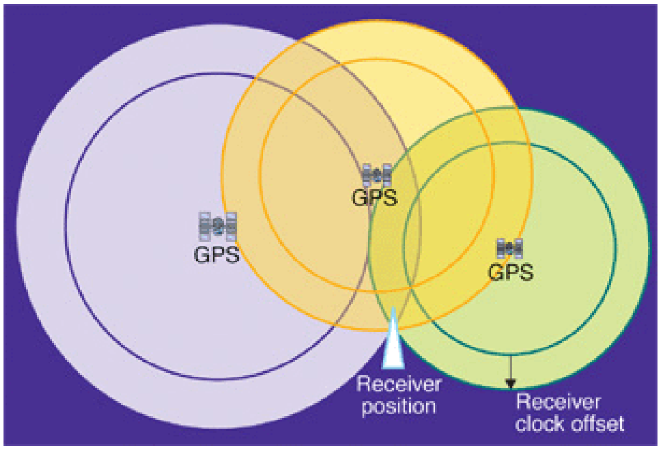


Figure 2.1: A GPS device determines its position based on the relative time delays of signals received from GPS satellites [1]

Using these techniques, a civilian GPS device can determine its absolute location anywhere on earth with an accuracy of 10-20m in most cases [1]. With more available satellite signals, a GPS device can get a more precise fix on its location. Thus, GPS is most effective where the view of the sky is unobstructed. GPS devices are less precise in dense urban areas where there are many buildings, since obstructions block or reflect radio signals, causing error in the location determination [7]. GPS devices also rarely work inside buildings, since they cannot receive reliable signals from the GPS satellites.

2.1.2 Applications

Only a few years after the completion of the GPS system in 1995, companies began introducing GPS devices aimed at consumers. In 1997, Garmin introduced the GPS III (see Figure 2.2), a device that could pinpoint the user and included an onboard database of roads and highways to help guide the user to his destination [8] [2]. Many other products from companies such as Magellan and TomTom hit the consumer market around the same time, allowing consumers to find their current location at any time. This was not only the first widely available method of geolocation, but it was also the first widely available location-aware application, since many of the GPS units included onboard maps.

Figure 2.2: The Garmin GPS III, a typical handheld consumer GPS device in the late 1990s [2]

Extending the idea of onboard maps, some units such as the Garmin StreetPilot (introduced in 1998) even enabled real-time navigation to guide the user turn-by-turn to a specific destination [8]. Most of these services, however, relied on onboard map databases, and were necessarily limited in the amount of data that they had available, which made them limited mostly to navigation. Location-aware applications that could help the user find specific services and attractions would come later when more interactive devices came to the scene.

Chapter 3

Defining Location-Based Technologies

Having developed an understanding of the technical and operational characteristics of previous-generation location-finding/location-based technologies, we can proceed to define the location-based technologies (LBTs) using previous generation technologies as a baseline. We look first at the some examples of LBTs, and then go on to identify defining characteristcs of modern LBTs.

3.1 Examples of LBTs

3.1.1 (Cellular Geolocation) Google Maps Mobile: My Location

While technically no different than previous cellular telephone geolocation approaches, Google Maps Mobile "My Location" feature marks one of the major transitions (in our mind) to modern LBTs. Released on November 28, 2007 [9], the "My Location" feature took advantage of the same technology used to provide some E-911 services – cell tower ID numbers. However, instead of relying on the cellular provider to perform the location lookup, Google transmitted the cell tower ID over the internet to its geolocation server to determine the user's location. Google also took advantage of users' telephones with built-in GPS, transmitting the (accurately-determined) GPS location to the server along with the user's current cell tower ID in order to further refine the Google geolocation database [10]. The location returned to the user's cell phone would then be transmitted to the Google Maps server to display the user's current location as a "magical blue circle" on a map of the current surrounding area.

This method of determining a user's cell phone location differs significantly from previous approaches. Instead of taking a passive approach or relying on an infrastructure provider to offer an authoritative answer, an interactive, data-driven approach is taken, exemplifying the "modern" LBT approach. The Google Maps Mobile software does offer the option to disable the "My Location" feature (although it is not emphasized)[9].

3.1.2 (Wireless LAN/WiFi Geolocation)

WiFi geolocation technologies take advantage of the popularity, ubiquity, and broad deployment of IEEE 802.11 wireless networks throughout urban areas, in homes, businesses, and institutions. WiFi geolocation services operate by capturing unique identifiers of wireless access points (APs), called basic service set identifiers (BSSIDs), broadcast by each AP within radio range. These identifiers are then looked up in a database that maps APs to geographic locations [11]. By examining the BSSID and intensity of each AP "seen" by a mobile device, the identification database can accurately determine the mobile device's location.

WiFi geolocation is becoming popular because of the advantages it offers over other geolocation methods. WiFi geolocation services offer a rapid time-to-first-fix (TTFF), the delay between turning on the service and determining your location, of less than a second compared to GPS' approximately thirty seconds. These services also operate much more effectively indoors and in urban areas, where cellular or GPS signals are poor or unavailable.

3.1.2.1 Skyhook Wireless: WiFi Positioning System

Skyhook Wireless is a major provider of WiFi-based geolocation services. Built on a large database of AP identifiers collected by "wardriving"[1] urban areas, Skyhook offers its WiFi Positioning System (WPS) to users in North America, Europe, and Asia [12]. The Skyhook WPS has one particularly interesting feature: it is "self-healing", or more accurately self-refining. When a location lookup is performed, any APs that are have moved or were never seen before are added to the location database to keep the system up-to-date[2]. WPS is also one component of Skyhook's recently released Hybrid Positioning System (XPS) that also incorporates GPS and cell tower information into geolocation requests to increase accuracy and availability[3]. In this case, cell tower and GPS location information may be used to

[1] scanning for AP identifiers and annotating them with the current geographic location determined using GPS, typically performed using a car

[2] `http://www.skyhookwireless.com/howitworks/wps.php`

[3] `http://www.skyhookwireless.com/howitworks/xps.php`

further refine the location database and the location query itself.

A number of popular devices and applications utilize Skyhook technology[4], including the Apple iPhone and iPod Touch, AOL Instant Messenger (via plug-in), Mozilla Geode, and Skyhook's own Loki service.

3.1.2.2 Google Gears Geolocation API

Google also offers a WiFi geolocation service through the Google Gears application programming interface (API) [13]. Little is known about this service, as Google has not disclosed its geolocation data sources [Ars], only stating that a "third-party provider" is being used[5]. An examination of the Google Android source repository provides some additional clues: a comment noting that "Service to communicate to the Google Location Server (GLS) via MASF server" and a server URL: `http://www.google.com/loc/m/api`[6].

The Google Geolocation API is believed to power the geolocation features of devices using the Google Android operating system, including the T-Mobile G1.

3.2 Technical Characteristics of New LBTs

As established in the previous chapter, early location-based technologies were passive and self-contained, having to rely only on their internal data stores and location received from external authorities such as GPS satellites or cellular towers to provide insight into their current location. Next generation LBTs, taking advantage of nearly ubiquitous wireless data service for mobile devices have deviated from this passive, self-contained approach in three major ways: they are interactive, data-intensive, and involve multiple providers per transaction.

3.2.1 Interactive

The first defining characteristic is that LBTs are *interactive*, relying on two-way data communication between the mobile device and a service provider to determine location or provide location-based services. While this may seem like a trivial distinction, and indeed it is not very interesting in and of itself, interactivity is the key enabling distinction of LBTs as compared to previous generations of technologies.

[4] `http://www.skyhookwireless.com/inaction/`

[5] http://code.google.com/p/gears/wiki/GeolocationAPI

[6] `http://android.git.kernel.org/?p=platform/frameworks/base.git;a=blob_plain;f=location/java/com/android/internal/location/LocationMasfClient.java;hb=master`

For example, when using Google Mobile maps on a mobile device to determine the location of the nearest gas station, your location and search query are trasmitted to the Google Maps server, which then responds with the name and coordinates of the gas station. This is in contrast to previous location-based technologies such as in-car navagation systems, where the names and coordinates of gas stations were maintained in an internal database inside the device, and then queried when the user asks for the nearest gas station.

This example illustrates one of the major functional benefits of interactive LBTs — that in providing the service they always query a central database, which should always contain the most up-to-date data.

3.2.2 Data-intensive

The second defining characteristic is that LBTs are *data-intensive*. What we mean is that instead of simply using the interactivity characteristic to replace an embedded local database with a centralized remote database, location data is used for additional purposes, such as improving a service's accuracy or providing additional contextual metadata for a service. Examples of some of these uses include:

- **Database population & training:** Skyhook and Google Gears Mobile geolocation systems rely on massive databases of geographic coordinates correlated to wireless access point cell tower identifiers. These services use data collected during location determination requests to populate and improve the accuracy of their databases.

- **Storage of data with associated location:** When capturing and uploading pictures to a photo sharing service such as Flickr, some LBTs devices such as the iPhone (or the EyeFi wireless location-aware camera memory card [7]) annotate the picture with the location where it was captured, for organization or future reference for the user and others.

- **Logging/retention of queries:** While it shouldn't come as a surprise, LBT service providers such as Google retain logs of search queries to train and improve their services [source]. While not drastically different from conventional search engine logging, in some cases these logs contain the geographic location included in the original query.[8]

[7] `http://www.eye.fi/`

[8] As discussed in an interview with Jane Horvath, Senior Privacy Counsel, Google Inc.

3.2.3 Multiple Providers Per Transaction

The third defining characteristic is that LBTs involve *multiple providers in completing a single LBT transaction*. In contrast to previous approaches at providing devices or services similar to modern LBTs where a single company provided a vertically integrated product (such as an in-car navigation system), modern LBTs rely on a variety of providers to satisfy a single user request. Presumably because of the cost and complexity of implementing some of the features required by LBTs, such as geolocation determination or geocoding (converting a civic address into a geodetic coordinate or vice versa), a layered model has arisen where LBT developers typically do not develop completely integrated solutions, but instead provide applications that then leverage other LBTs such as the Google Maps API or the Skyhook Wireless geolocation service. Descriptions and examples of each of the relevant layers is given in table 3.1.

Table 3.1: Description of layers in LBTs

Layer	**Function**	**Examples**
Application	Providing or determining application-specific data	Urbanspoon
Mapping	Rendering a user's current location and other points of interest on a map	Google Maps API, Microsoft Live Maps
Geocoding	Converting a civic address (street address) into geographic coordinates (latitude & longitude)[9] or vice versa	Yahoo Maps API, `geocoder.us`, Google Maps API
Geolocation	Determining a user's location based on observed information such as wireless access point or cell tower identifiers	Skyhook, Google Gears API

An Illustrative Example In developing an application using this layered approach, a developer passes information to the provider in the layer they are interested in obtaining service from, and then pass the result or other data to a provider in the next layer of interest, until the desired transaction is complete. For example, a hypothetical location-based application on the Apple iPhone to display real-time traffic congestion information could involve communicating location and search query information with the following parties to complete a single transaction:

- **Geolocation provider (Skyhook)**, to determine user's geographic coordinates using WiFi positioning

- **Geocoding provider (Google Gears)**, to convert user's geographic coordinates (latitude & longitude) to a civic address (street name)
- **The application provider**, to obtain congestion data for the specified street name
- **Map Provider (Google Maps)**, to request and display a map of the user's location with overlaid traffic congestion information

and of course all of this data is transmitted across the cellular provider's network.

3.3 User Perceptions of New LBTs

In understanding the nature of new LBTs, it is also valuable to consider the perceptions of the users of these technologies. While GPS-based handheld receivers and in-car navigation systems have been widely available for some time, the shift to having location-determining devices embedded in cellular phones has been a recent change. Beginning with GPS-equipped cell phones and Google's My Location feature, the locating and navigation functionality that used to only be used while hunting or driving is now increasingly in your pocket, and always available.

While we did not conduct research into user perception of new LBTs, we posit, based on our own personal experiences with iPhone and G1 users, news reports[10] and LBT marketing[11], that the user experience is affected in the following ways:

- New location-based technologies offer **new functionality** that is, to the credit of LBT developers, often well-packaged and very easy to use compared to previous product offerings. For example, some particularly well-written iPhone applications could likely be claimed to be "indistinguishable from magic" (to quote Arthur C. Clarke).
- Also to the credit of LBT developers, these new technologies offer very **useful and convenient functionality**. From finding a nearby coffee shop to simply finding your location quickly and more accurately (see figure 4.1), these technologies can provide significant benefit to users.
- Finally, while perhaps not directly perceived by most users, these new technologies are making concepts that we associate primarily with desktop computing — social

[10] http://www.forbes.com/technology/personaltech/2008/06/15/iphone-apps-appeal-tech-wireless08-cx_bc_0616apps.html

[11] http://news.bbc.co.uk/2/hi/technology/6246063.stm

networking, searching, or even advertising — **ubiquitous** even while away from the computer. While this is perhaps a subtle or vague point, it is important in that it changes user expecations, including expectation of privacy.

3.4 The Privacy Tension in LBTs

These technical characteristics of new location-based technologies raise interesting and potentially concerning implications for user privacy: who has information about my location, and how are they using it? It is very easy to "put on a tin-foil hat" and proclaim that because location-based technologies involve communicating your current location — arguably private information — to third parties, your privacy is at risk. However, this (sadly) is, or should be, *obvious*. There are inherent privacy risks in communicating any information we deem private to third parties, and this is no different for location-based technologies. What *is* more subtle is that because of the defining technical characteristics of LBTs — most notably the multiple providers per transaction characteristic — it is extremely difficult for users to understand or to manage their privacy risks when using modern location-based technologies.

In tension with the complex technological nature of LBTs is the user perceptions of LBTs and their desire to participate in technical innovation or enjoy convenient new technologies. Due to the difficulty of understanding the privacy risks of LBTs, users may unknowingly find themselves in a situation of asymmetric information with respect to the privacy risks they face. Users can only make appropriate decisions to manage their privacy risks if they are fully informed, which may lead them to accepting privacy risks that are incompatible with their personal privacy values. Unlike more mature technologies and platforms that inform users and provide choice regarding privacy, both the nature of location-based technologies and the implementations of current popular services make it difficult to understand the privacy risks associated with using these services, and provide few options, aside from abstaining, to manage privacy risks.

Chapter 4

Privacy in Location-Based Technologies

With a greater understanding of the subtle privacy issues and risks, both in terms of absolute risks as well as the problems with managing risk, we analyze some existing location-based technologies (LBTs) to illustrate both the risks and the difficult of managing risks associated with today's location-based technologies.

4.1 Privacy Analysis of Location-based devices

Mobile devices, while perhaps least involved in actually providing location-based technologies, serve as the platform for the use of location-based applications and services, and so they can play an important role in determining the privacy risks associated with using a location-based technology. In particular, highly integrated platforms such as the Apple iPhone are particularly important in determining the privacy risks and approach to privacy management platform.

4.1.1 Apple iPhone

The Apple iPhone continues to be one of the world's most popular smartphones, having sold over thirteen million units since its release just over one year ago [14]. It plays a particularly important role in the context of location-based technology, as Apple has placed an emphasis on encouraging location-based applications, providing developers with the Core Location API to facilitate location-based development. According to Skyhook Wireless, there are over 600 location-based applications in the iTunes App Store as of early December 2008

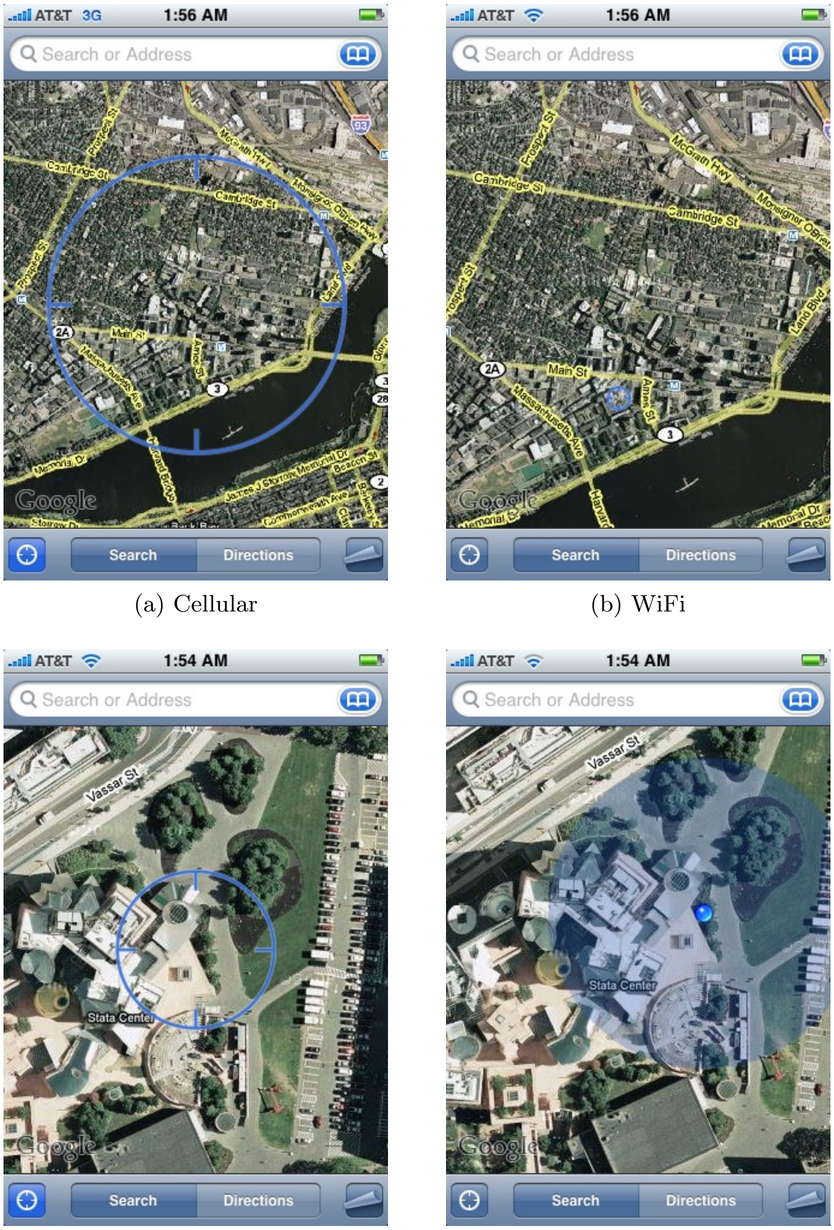
[15]. Apple readily provides the Core Location API to developers to foster application creation, but Apple retains complete control over the applications available to customers through the App Store. Apple can (and often does) decide to reject apps for "duplicating functionality" or other offenses against Apple's iPhone developer agreement [16]. All queries for core iPhone services are routed through Apple's servers as well. Thus, Apple holds almost complete control over the privacy effects of the iPhone platform.

4.1.1.1 iPhone Location Framework

The Core Location API gives applications on the iPhone access to the hybrid positioning system licensed from Skyhook Wireless to determine the phone's position. In the original iPhone, the location was determined using a combination of cell-tower and WiFi access point location methods. In the iPhone 3G, though, an assisted GPS unit was added to further enhance the accuracy and availability of the location determination service. All experiments in this section were performed using an iPhone 3G, and this paper in general assumes an iPhone 3G as the norm.

The hybrid location determination can be extremely accurate, as shown in Figure 4.1. Using cellular-only location (achieved by turning off WiFi and standing far enough away from a window so as to not receive a GPS fix), the location determination algorithm can only pinpoint the phone's location to the granularity of a section of a city. Figure 4.1 (a) includes the majority of MIT's campus, as well as a large portion of the surrounding Cambridge area. Turning on WiFi significantly narrows the location region (as shown in (b) and (c) of Figure 4.1), and the phone can now determine the exact building it is within on MIT's campus and even the *section* of the building. When adding a GPS signal to this determination (accomplished in Figure 4.1 (d) by holding the phone out the window of the same room), the location can be further refined.

Any application on the iPhone can ask the Core Location API to make a location determination, and the Core Location engine will use any available methods (cellular, WiFi, and/or GPS) to make an accurate location determination. The first time an application attempts to use the Core Location API, the user is presented with a popup window confirming that he does indeed want the application to access the phone's location information (as seen in Figure 4.2). There is a menu item in the phone settings to reset these location warnings and clear all previously stored allowances for applications.



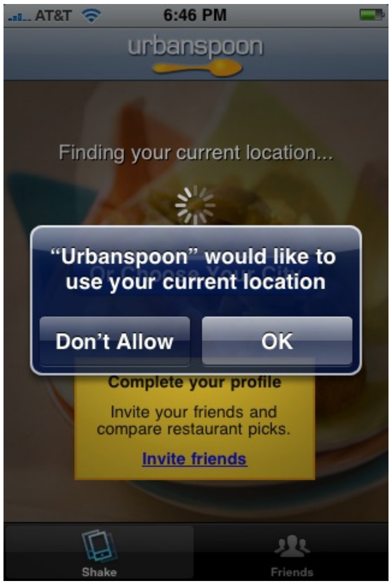
(a) Cellular

(b) WiFi

(c) WiFi (re-scaled)

(d) GPS

Figure 4.1: Apple iPhone geolocation of MIT 32-G882 using various technologies



Finding your current location...

“Urbanspoon” would like to use your current location

Complete your profile

Invite your friends and compare restaurant picks.

Figure 4.2: The Core Location API prompts the user for permission when an application first uses Core Location

4.1.1.2 Privacy Policies

On the iPhone, the “Legal” section can be found by navigating through a rat’s nest of menus: “Settings” to “General” to “About” to “Legal” (all the way at the bottom). Once in the Legal section, there are pages upon pages of copyright notices, but little mention of user privacy. There is, upon close inspection, a small “Privacy Policies” section buried among the legalese which includes links to two privacy policies, and the full text of one more policy that can be found replicated online:

- Apple’s Privacy Policy - `http://www.apple.com/legal/privacy/`
- YouTube’s Privacy Notice - `http://www.youtube.com/t/privacy`
- Google Maps Privacy Notice - `http://web.archive.org/web/20070701102817/http://www.google.com/mobile/privacy.html`

Upon close examination of Apple’s privacy policy, it quickly becomes apparent that this policy was not designed with location information in mind. The closest the policy ever comes to acknowledging location is when it mentions that Apple “also collect[s] information for market research purposes such as your occupation and where you use your computer to gain a better understanding of our customers and thus provide more valuable service." This, however, seems to clearly imply that "where you use your computer" is merely a determination of the environment in which a user is most active with his computer (e.g. home, work, or school), since there is a very similarly worded question in the setup and registration process for Apple computers. It does not appear to cover specific physical location as determined by Core Location on the iPhone. The policy generally states that Apple protects users' information, and may share it with partners from time to time.

The YouTube privacy policy is not of interest at all for the scope of location services, since (at the time of writing) YouTube does not offer any location-based content or store location tags of any kind.

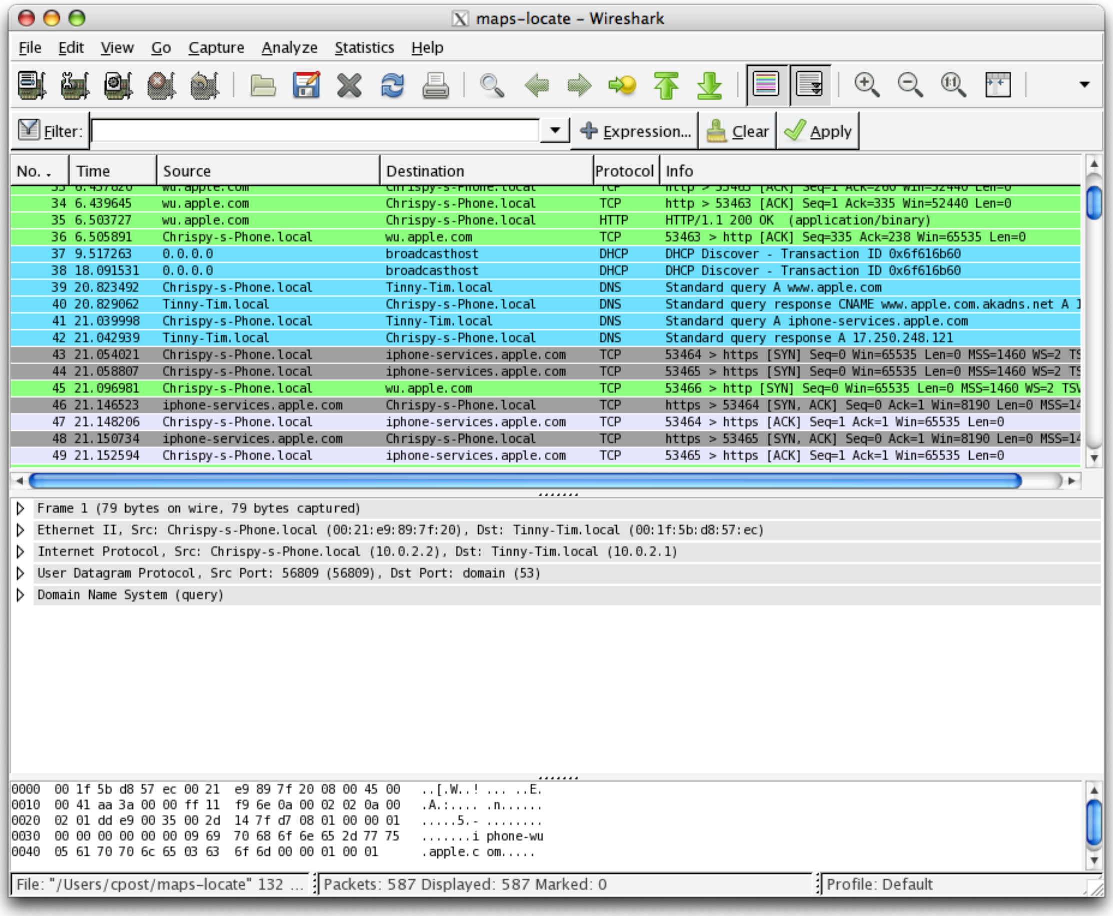
The Google Maps Privacy notice is the only of these privacy policies to mention location information explicitly: "If you use location-based products and services, such as Google Maps for mobile, you may be sending us location information. This information may reveal your actual location, such as GPS data, or it may not, such as when you submit a partial address to look at a map of the area." The policy also acknowledges that this information may be stored in order to personalize services to the user, and may be released in an aggregate, anonymous fashion to third parties. The specific queries, though, are not forwarded outside of Google except where absolutely necessary for business purposes.

None of these policies, however, say exactly where location queries are actually sent. Skyhook Wireless provides the location determination algorithm for the iPhone; are the location queries sent to Skyhook? Are they sent to Apple? The specifics of where personal location information travels is largely lacking in these policy descriptions. In order to determine where the location queries are *actually* sent, some experimentation is necessary.

4.1.1.3 What's Really Behind Core Location

In order to determine the actual destination of location queries and other communications to or from the iPhone, all of the network traffic for the iPhone was fed through a computer to monitor the network conversation. When an iPhone is associated to a WiFi network, it favors that network link over the cellular data connection, so all the traffic travels over the WiFi link. A laptop was connected to a wired network with connectivity to the internet, and this connection was shared over the WiFi connection. The iPhone was connected to the computer's WiFi connection, so all it's traffic passed through the computer. Wireshark (available at `http://www.wireshark.org/`), a popular network traffic capture and analyzation tool was then run on the computer to capture all the network traffic coming to or from the iPhone.

First, the location warnings were reset in the phone settings. Google Maps was opened. At this point, there was substantial unencrypted traffic to the wu.apple.com server. Inspection of this network traffic revealed no location information being transmitted, however. The contents of the traffic to texttttwu.apple.com revealed the fact that the phone had opened Google Maps and the application and firmware version numbers of the phone. As soon as the location privacy dialogue was confirmed, however, a secured (SSL) connection to `iphone-services.apple.com` was opened (just before packet number 39 in Figure 4.3), presumably to obtain a location fix by using the cellular and WiFi databases licensed to Apple by Skyhook Wireless.



No.	Time	Source	Destination	Protocol	Info
34	6.439645	wu.apple.com	Chrispy-s-Phone.local	TCP	http > 53463 [ACK] Seq=1 Ack=335 Win=52440 Len=0
35	6.503727	wu.apple.com	Chrispy-s-Phone.local	HTTP	HTTP/1.1 200 OK (application/binary)
36	6.505891	Chrispy-s-Phone.local	wu.apple.com	TCP	53463 > http [ACK] Seq=335 Ack=238 Win=65535 Len=0
37	9.517263	0.0.0.0	broadcasthost	DHCP	DHCP Discover - Transaction ID 0x6f616b60
38	18.091531	0.0.0.0	broadcasthost	DHCP	DHCP Discover - Transaction ID 0x6f616b60
39	20.823492	Chrispy-s-Phone.local	Tinny-Tim.local	DNS	Standard query A www.apple.com
40	20.829062	Tinny-Tim.local	Chrispy-s-Phone.local	DNS	Standard query response CNAME www.apple.com.akadns.net A 1
41	21.039998	Chrispy-s-Phone.local	Tinny-Tim.local	DNS	Standard query A iphone-services.apple.com
42	21.042939	Tinny-Tim.local	Chrispy-s-Phone.local	DNS	Standard query response A 17.250.248.121
43	21.054021	Chrispy-s-Phone.local	iphone-services.apple.com	TCP	53464 > https [SYN] Seq=0 Win=65535 Len=0 MSS=1460 WS=2 TS
44	21.058807	Chrispy-s-Phone.local	iphone-services.apple.com	TCP	53465 > https [SYN] Seq=0 Win=65535 Len=0 MSS=1460 WS=2 TS
45	21.096981	Chrispy-s-Phone.local	wu.apple.com	TCP	53466 > http [SYN] Seq=0 Win=65535 Len=0 MSS=1460 WS=2 TSV
46	21.146523	iphone-services.apple.com	Chrispy-s-Phone.local	TCP	https > 53464 [SYN, ACK] Seq=0 Ack=1 Win=8190 Len=0 MSS=14
47	21.148206	Chrispy-s-Phone.local	iphone-services.apple.com	TCP	53464 > https [ACK] Seq=1 Ack=1 Win=65535 Len=0
48	21.150734	iphone-services.apple.com	Chrispy-s-Phone.local	TCP	https > 53465 [SYN, ACK] Seq=0 Ack=1 Win=8190 Len=0 MSS=14
49	21.152594	Chrispy-s-Phone.local	iphone-services.apple.com	TCP	53465 > https [ACK] Seq=1 Ack=1 Win=65535 Len=0

```
▷ Frame 1 (79 bytes on wire, 79 bytes captured)
▷ Ethernet II, Src: Chrispy-s-Phone.local (00:21:e9:89:7f:20), Dst: Tinny-Tim.local (00:1f:5b:d8:57:ec)
▷ Internet Protocol, Src: Chrispy-s-Phone.local (10.0.2.2), Dst: Tinny-Tim.local (10.0.2.1)
▷ User Datagram Protocol, Src Port: 56809 (56809), Dst Port: domain (53)
▷ Domain Name System (query)
```

```
0000  00 1f 5b d8 57 ec 00 21  e9 89 7f 20 08 00 45 00   ..[.W..! ... ..E.
0010  00 41 aa 3a 00 00 ff 11  f9 6e 0a 00 02 02 0a 00   .A.:.... .n......
0020  02 01 dd e9 00 35 00 2d  14 7f d7 08 01 00 00 01   .....5.- ........
0030  00 00 00 00 00 00 09 69  70 68 6f 6e 65 2d 77 75   .......i phone-wu
0040  05 61 70 70 6c 65 03 63  6f 6d 00 00 01 00 01      .apple.c om.....
```

File: "/Users/cpost/maps-locate" 132 ... | Packets: 587 Displayed: 587 Marked: 0 | Profile: Default

Figure 4.3: Network traffic during a Google Maps location fix. Location confirmation dialogue box was confirmed just before packet number 39

The connection to the geolocation database is encrypted, which makes the contents of the communication impossible to read. While this allows little information to be gleaned about the nature of the location determination transaction, it does help preserve the privacy of the user's location by ensuring that a third party listening to the network connection could not glean a users' location solely from the location determination mechanisms. What is notable, however, is that the connection happens to a server in Apple's control, not Skyhook's. This further enforces the paradigm of Apple's complete control over the iPhone platform, making Apple responsible for the majority of the privacy-related concerns on the iPhone.

4.1.1.4 How Do Third Party Apps Fit?

The Core Location API itself preserves the integrity of users' location data as it travels across the network (either cellular or WiFi) by using SSL encryption. Even so, there are hundreds of location-aware applications available for the iPhone today. Do these applications follow the same standards?

To find out, a selection of ten iPhone applications were tested using the methodology from the previous section to listen to the network traffic during the use of each application. Using the "Top 10 Apps" listing for various categories on `http://www.apple.com/iphone/appstore/`, ten location-based applications were selected for testing:

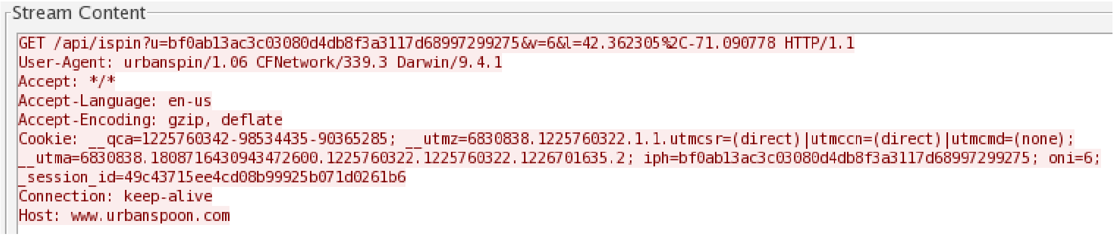
- Urbanspoon - An application to help user find restaurants in their vicinity.
- Google Maps - A built-in application for the iPhone (the only one considered here). It can plot the user's current location on the map, search for nearby services, and plot driving directions.
- Google Earth - Similar capabilities to Google Maps, but shows the earth's surface in 3D and allows the user to fly around the landscape.
- The Weather Channel - Provides weather information for the user's local area.
- AroundMe - Provides an easy search for gas, food, movies, and other services near the user, sorted by category.
- Bank of America - Provides an interface to electronic banking services, but also allows the user to look up the location of nearby ATMs and banking centers.
- WeatherBug - Similar to The Weather Channel, this application provides weather info for the user's area.

- YP Mobile - A mobile interface from YellowPages.com where users can look up businesses and services near their current location.

- Yelp - Allows users to find restaurants, bars, gas stations, and other services near their current location.

- Where - Another interface to help users find locations of interest near their current location.

Urbanspoon was the first application tested. After allowing the application to use the Core Location API, the phone connected to `iphone-services.apple.com` over an SLL connection to use the cellular/WiFi location database. When the "shake" button in Urbanspoon was pressed (this picks a restaurant in the nearby area for the user), the application made an *un*secured connection to `www.urbanspoon.com` in order to find restaurants (the request is shown in Figure 4.4). In the "`HTTP GET`" request, there are many parameters passed to `www.urbanspoon.com`, including "`l=42.362305%2C-71.090778`", which reveals GPS coordinates 42.362305, -71.090778 as the precise position of the iPhone as determined by the Core Location API. As this data is transmitted in plaintext, any party listening to the network at any point directly between the iPhone and `www.urbanspoon.com` can determine the iPhone's location through simple inspection of the network traffic. (It took less than a minute of inspection to find the phone's coordinates from the network traffic in this case.) Also, since the GPS coordinates of the iPhone were transmitted through a "`GET`" request, they are visible in the URL of the transaction, not in the body. Many web server systems consider the URL to be transactional data and log this information, where they would not log the body of the transaction. Thus, within the realm of internet information, the precise location of the iPhone in a query to Urbanspoon is in a class of information with little to no protection mechanisms.



```
GET /api/ispin?u=bf0ab13ac3c03080d4db8f3a3117d68997299275&w=6&l=42.362305%2C-71.090778 HTTP/1.1
User-Agent: urbanspin/1.06 CFNetwork/339.3 Darwin/9.4.1
Accept: */*
Accept-Language: en-us
Accept-Encoding: gzip, deflate
Cookie: __qca=1225760342-98534435-90365285; __utmz=6830838.1225760322.1.1.utmcsr=(direct)|utmccn=(direct)|utmcmd=(none);
__utma=6830838.1808716430943472600.1225760322.1225760322.1226701635.2; iph=bf0ab13ac3c03080d4db8f3a3117d68997299275; oni=6;
_session_id=49c43715ee4cd08b99925b071d0261b6
Connection: keep-alive
Host: www.urbanspoon.com
```

Figure 4.4: The unsecured HTTP request to `www.urbanspoon.com` to find restaurants in the user's vicinity; the request reveals the user's precise coordinates

The other nine application were tested in the same manner as Urbanspoon. For The

Weather Channel, AroundMe, Bank of America, WeatherBug, YP Mobile, Yelp, and Where, precise GPS coordinates could be easily extracted through inspection of the network traffic. For Google Earth, coordinates could not be easily extracted from the communication due to the large amount of data received from Google's servers, but the communication between the phone and Google's servers was not encrypted. This means it should be possible to devise a method for extracting the location from the network traffic, but the implementation of this method is beyond the scope of this paper. For Google Maps, the transmission of map data appeared to come from `iphone-maps.apple.com`, but the connection was encrypted though SSL, so the actual contents of the communication is unknown. In summary, 8 of 10 application tested transmitted precise location information obviously over an unsecure connection, 1 did not use encryption, but made the location information hard to find, and only 1 (the only built-in application tested) made an effort to secure location data through encryption.

Location privacy with third party applications is not covered in Apple's privacy policies. It is, however, mentioned in the iPhone SDK agreement that developers must agree to in order to have their app distributed through the App Store. For applications that "collect, transmit, maintain, process, share, disclose or otherwise use a user's personal information," the application must adhere to applicable privacy laws, but the policy does not explicitly hold third parties to any higher standard [17]. Furthermore, any applications that use the Core Location API "must notify and obtain consent from the individual" before location data is used. The agreement, does not, however, specify how this information can be communicated to third party servers, leaving the door open for applications to communicate private location data unencrypted.

4.1.1.5 Conclusion

Apple exerts control over almost all of the aspects of the iPhone platform; Apple has power over the privacy controls associated with the device. Even so, the privacy policies associated with the iPhone are deeply buried in menus, and are vague with respect to user location information. The technical implementation of the Core Location API protects the user's location when a location determination is made using encryption, but the same standards are not applied to third party applications. The overwhelming majority of these third party applications do not encrypt user location data as it travels across the network, leaving the user's location exposed to anyone who can intercept the network traffic. Users cannot only trust Apple with the privacy of their location data; they are required to trust that various third party services apply proper privacy controls to their personal data as well.

Apple attempts to address this in the iPhone SDK Agreement, but does not hold third parties even to the same standard as Apple's Privacy Policy, which itself is vague with respect to location information. Apple could easily implement stronger controls on the use and transmission of user location data, and hold third party developers to standards more strongly. Apple could also allow the user more control over the use of location data. However, as it currently stands, it is difficult for iPhone users to accurately determine the privacy risks associated with using the device, or to manage these risks in accordance with their personal privacy values.

4.1.2 T-Mobile G1

The T-Mobile G1, released in the United States on October 22, 2008 for a price of $179 (with a 2-year contract) was the first commercially available phone to use the Android operating environment [18]. Android, a product of Google and the Open Handset Alliance, aims to reduce the costs associated with development on mobile device platforms, give developers an open platform that will allow them to collaboratively develop software, and give customers a unique and cutting-edge technology experience [19]. Android and the G1 also place a high importance on location-based services. More than half of the top 10 applications in the Android Developer Contest were location-based [20]. There is no one entity that controls all aspects of the G1 phone, which provides unique challenges for user privacy.

4.1.2.1 G1 Location Framework

The android.location package provides location services for the G1 phone [21]. Unlike Core Location on the iPhone, which can only allow an application to ask the API for the device's current location using the method Core Location deems most accurate, android.location enables a flexible interface from the application's view:

- The Android location framework allows for multiple "location providers," each of which can give an estimate about the device's position. On the G1, there is a GPS location provider that uses the internal GPS unit, and a location provider that uses the cellular tower and WiFi access point signal strengths to reference an external database to determine location.

- The LocationManager class allows the application to view a list of location providers, choose a location provider, and get a location fix from any available location provider.

If the application wants to take into account the position information from multiple providers to determine a more accurate location, it must do so on its own.

- LocationManager also allows an application to request periodic updates on location position, even when the application is not active. This is very different from the iPhone platform, where applications cannot run in the background. With Android, an application can poll for locations updates even while is is not active in the foreground. This is a useful feature, but allows for more potential misuse of location determination by application developers.
- Another ability that LocationManager provides for applications is a method for an application to register an alert for when the device is within a specified radius to a given location. This allows the application to trigger events when certain destinations are reaches without periodically polling the location provider.

On Android, per-application location privacy is managed by prompting the user at install time about whether to allow the application to use location information. This permission cannot be revoked without uninstalling the application or disabling location services on a global basis. As with the iPhone, this lack of per-app control at any time is a serious shortcoming for users who would like more fine-grained control of their privacy.

4.1.2.2 Privacy Policies

On the phone, the G1 Privacy Policy [22] can be accessed through the legal section of the settings menus, in a similar fashion to the iPhone. It is, however, not buried quite as deeply as the iPhone's policies. While certainly more comprehensive and informative than the iPhone Privacy Policy, the G1 Privacy Policy is not without problems.

Up front, the G1 privacy policy addresses the fact that third party applications are not covered under the G1 Privacy Policy, and the document encourages users to seek out the privacy policies of those individual applications. This is a step in the right direction, since it encourages awareness, user involvement, and acknowledges the distributed nature of modern location-based technologies. Ideally, however, users would have a simple way of looking at all the privacy policies concerning their personal data.

The policy only briefly mentions location information as well. The G1 “may send [Google] location information (for example, Cell ID or GPS information) that is not associated with your account” [22]. This is a clear, easily understandable statement that informs the user that location data is being transmitted to Google. It also states that location data is not paired with a Google account, giving the user some level of anonymity. In practice, however, it would be easy for Google to infer the account associated with a particular request, since the IP address can be tied to other requests that *do* make use of a user's Google account.

It is clear that Google obtains a user's location data during a location determination transaction. The privacy policy states that Google will not share personal information with third parties except to conduct necessary business tasks such as billing, or in aggregated form to provide general usage statistics that do not identify individual users.

4.1.2.3 Third-Party Applications

As with the iPhone, location-based applications have become popular in the short time the G1 has been available. Many of the winning application from the Android Developer Challenge were location-based, and provide similar services to some of the iPhone applications previously mentioned [20]. Many applications allow users to find services and points of interest in their vicinity. Some, like Wertago, bring location to social networking, allowing users to find out how many people are gathering at party spots around the city and broadcast their own locations to friends. Most of these applications have similar risks to their iPhone counterparts. Users must trust the third party service to protect their location data, and must also be concerned with precise location data being transmitted unencrypted across the network.

There is, however, a very different threat model applicable with the G1's location framework. Third party applications can periodically query the location service, even when they are not active. Locale is an application that can change user settings based on the location of the device, as well as other factors [20]. In order for applications like this to operate, the Android location framework must determine the user's location even when the user is not actively using a location-aware application. This "always-on" nature causes location queries (in the case of cellular/WiFi based location determination) to be executed much more often than if they were only executed when the user had a location-aware application open. This might allow closer tracking of an individual through various channels, such as network traffic sniffing or compiled third party data.

4.1.2.4 Conclusion

The G1, though very new at the time of writing, has proved to be very popular and useful. It, like the iPhone, puts a strong emphasis on location-aware applications, and many developers have created applications in this category. Google takes a step in the right direction by acknowledging that location data is transmitted using the phone. Google also acknowledges that third parties might have access to personal location data, but it does not provide a user-friendly way to determine exactly what risks are present with an individual third party. The nature of LocationManager also allows applications to take more control of the location interface, putting less burden on the operating system for determining privacy and more on the application. This causes a complicated web of trust for users, and it is difficult to determine what parties are involved in handling personal location data for any given location service. Furthermore, there is a lack of easily-definable and fine-grained privacy controls for which applications and services can make use of location data. Again, as with the iPhone, it is difficult for users of the G1 to understand or control the privacy risks associated with the device.

4.2 Location Determination Services

The new class of location-based technology is largely founded upon the ability to determine location without relying solely on a passive service such as GPS. While GPS does play an important role in both the iPhone and the G1's location infrastructure, the primary reason why these devices are so useful (especially in urban areas where GPS does not work effectively) is due to the cellular and WiFi geolocation methods. The nature of these methods makes a database of cellular towers and WiFi access points necessary to perform a location determination. This database cannot easily be stored on the device itself, as it would occupy vast amounts of storage and would quickly become unreliable due to changes in the WiFi landscape. Thus, services where a mobile device can query a centralized database are necessary.

4.2.1 Skyhook Wireless

Skyhook Wireless was founded in 2003 to meet the rising demand for location determination technologies [23]. GPS and cellular tower triangulation proved inadequate for mobile application because of slow determination time and unreliability in urban environments. Because of the pervasive nature of WiFi access points by that time, Skyhook was able to build a database of access points and develop location determination technology that looked at the unique addresses (BSSIDs) of access points a device could see to make a determination. Skyhook Wireless also provides the location determination service for the Apple iPhone [15].

4.2.1.1 Architecture Overview

In order for the WiFi positioning system (WPS) to make a location determination, there must be a large database of access points available. Skyhook first develops coverage for a particular area by "wardriving" the streets of a city [12]. While driving around every street available with WiFi scanning equipment and a GPS location device, data with the GPS coordinates and the access points available at those coordinates are recorded. These data can then be used to develop a location estimate for each unique WiFi access point.

With the XPS hybrid positioning system, mobile devices also provide training data to help update the WiFi database. For example, if a mobile device can see five access points with known locations in the database, a location determination can be made. If the mobile device also sees a sixth, unknown access point as well, the database can make an estimate for that new access point's position and begin using it in the database. Over time, this device-based training helps keep the database up to date.

Because the database is a critical piece of Skyhook's intellectual property, it is certainly not available for full scrutiny by the public. Devices do not even have full access to the database, as they send "what they see" to the location determination service, and the service sends back an estimate of the device's position. There is, however, a function in Skyhook's developer interface that allows anyone to query the location of a specific BSSID [24]. This interface is undocumented by Skyhook, but makes the estimated location of any access point in the database readily available to anyone with some knowledge of computers. As mentioned in Chapter II where Bob tracked his ex-wife's (more precisely, her access point's) location across the country, this hole can be exploited to enable tracking of access points once their location has been updated in the database. This could potentially compromise the privacy of users who aren't even participating in location-based technology. Furthermore, Skyhook provides no method to remove or blacklist an access point from the database, so it is nearly impossible for a user to mitigate this threat.

4.2.1.2 Privacy Policy

Skyhook's Privacy Policy is overall very comprehensive and detailed. It is very explicit that it does not collect any other identifying information about the user other than the data needed to determine location [25]. Skyhook even goes so far as to state that "Skyhook Wireless CAN NOT, DOES NOT and WILL NOT track your location" in the policy. Each installation of the mobile device client does have a unique identifier that is sent to Skyhook, so it would be possible to track a specific device across multiple queries if Skyhook modified their technical architecture, but as it is specified in the privacy policy, the unique identifier for each user is not stored in the transaction log, so it would be impossible to retroactively track a user. Furthermore, the data from these anonymized logs is removed from the production servers after 72 hours and stored in encrypted form on a backup system. Skyhook also uses SSL encryption to protect all communications with the wireless positioning clients.

4.2.1.3 Conclusion

Skyhook makes many strong assertions in its privacy policy that value and protect users' location information. Many of these could provide a model for other companies who deal with location information to assure users of the security of their personal location data. Skyhook also uses strong encryption not only on the communications with mobile devices, but in their long-term storage database of location queries as well. Despite the excellent protection of user privacy from the policy side, however, Skyhook has one glaring flaw in its technical infrastructure: the ability for anyone to determine the location of a particular access point in the database. This flaw could be easily fixed, however, since is is not needed for the normal operation of the service.

Chapter 5

Improving Privacy

As shown in the previous chapters, there are significant gaps in the legal, corporate policy, and technical aspects of location-based technology (LBT). LBTs have successfully grown to be ubiquitous among consumers, and they will only continue to grow more rapidly in years to come. In order to allow these technologies to continue growing without confusing users and exposing them to privacy risks, the LBT industry as well as the U.S. legislative branch need to ensure that these privacy concerns are addressed. In this chapter, we will propose solutions to the problems outlines in previous chapters to mitigate the privacy concerns inherent in location-based technology.

In order to maintain the usefulness of LBT services while still minimizing privacy issues, these solutions aid in preserving the following key values that are essential to LBT:

- Governments and LBT providers must realize location data reveals inherently private and valuable information about users: Any one particular location query may seem relatively harmless by itself, but the frequency at which these location queries happen (due to the widespread use of LBTs) means a very specific picture of a user's habits can be painted. Even if the data is anonymized, the location data would commonly be enough to determine where a user lives and works, greatly narrowing the list of possible real identities behind the "anonymous" identifier. People value the privacy of their location, and government regulation and corporate policy should reflect this.

- Users must be informed: The nature of modern LBT causes users' private location data to be shuffled through various companies and services in order for the technology to operate. While the distribution of this information is inevitable, users should be able to determine where exactly their location information is traveling and what specific privacy policies apply to this data. Users should also know if a specific company is retaining or releasing their data and in exactly what form this happens.

- Users must be given fine-grained choice: Even on the same device platform, different location-based services might expose the user to different risks, since each application can be served by a different provider. Users should be able to decide which individuals applications they would like to share their location information with. They should also be able to decide the granularity of information to be shared with various locations. For example, a user should be able to specify that precise coordinates be shared with some applications and only the city where he is be shared with other applications.

5.1 Legal Approaches

5.1.1 Private Dataholder Accountability

Given that constitutional protections of privacy such as the *Fourth Amendment to the United States Constitution* only protects a citizen's privacy from agents of the government, protecting an individual's privacy from invasion by other non-government parties must come from another source. In our research, we identified three primary sources of privacy protection governing private parties:

- Overarching laws, such as the European Union Directive on the Protection of Personal Data[1], which requires the state to regulate the processing of personal information by any entity

- Industry self-regulation, such as the *Best Practices and Guidelines for Location-Based Services*[26] proposed by CTIA – The Wireless Association, which provides a set of guidelines for industry members to follow, and permits them to "self-certify" that they are following the guidelines.

- Individual self-regulation, based on customer/market demands, etc. such as that espoused by Google: "We make privacy a priority because our business depends on it. In fact, if our users are uncomfortable with how we manage their personal information, they are only one click away from switching to a competitor's services."[27]

All of these approaches have different focuses, though the examples given above all contain some elements regarding informing the user, providing choice, and controlling data.

[1] http://ec.europa.eu/justice_home/fsj/privacy/index_en.htm

We propose a different approach, one that is not new, but that is particularly well-suited to this type of problem. This approach is that of *information accountability*, proposed by Weitzner et al. in a paper by the same name[28].

5.1.1.1 The Need for Accountability Instead of Control or Opt-out

As we have discussed previously, it is unrealistic to attempt to stem or control data collection as a means of protecting user privacy. LBTs provide convenient and useful services, and the architectural approaches to implementing LBTs means that data is retained by and shared between multiple providers to complete a transaction. The effects of trying to control information or insisting on opting-out as the only privacy control measure would surely hamper LBTs, preventing innovation and reducing the convenience of using these services. However, without any protection of LBT data, the potential for abuse could be large, and the consequences for personal privacy from abuse of said data would be significant[2]. Controls are clearly necessary, but we must look beyond conventional concepts of restricting access, as suggested by proponents of information accountability[28].

In their paper, Weitzner et al. look to the United States' Fair Credit Reporting Act (FCRA) as an example of an information accountability legal framework in action. The rationale behind the FCRA reflects some of the needs for protecting LBT location information, illustrating the potential benefits of a FCRA-like legislative approach to holding private data holders accountable to the people whose data they collect.

The FCRA was enacted in 1970 in response to complaints of inappropriate data collection and use by credit agencies [29]. Credit reporting was a useful service provided by credit agencies, but the data held by these agencies was sometimes disclosed improperly or used for inappropriate types of decision-making. The FCRA lawmakers, presumably recognizing the benefits of credit reporting and also that credit reporting relied on (accurate and properly-gathered) data, did not attempt to control or limit handling of consumer credit data. Instead credit agencies were made responsible for using this data for authorized purposes only, with financial penalties and jail time to dissuade those considering non-compliance. In terms of efficacy, [28] notes, "The success of this accountability regime for the past 40 years over a very large set of data — credit reports on nearly every adult in the U.S. — makes it a worthy model for considering policy compliance in other large systems."

[2] Chapter II illustrates the potential consequences of the subtle privacy problems associated with LBT. Blatant abuse/improper access to LBT data could allow the tracking or profiling of an individual without their knowledge or consent.

5.1.1.2 Principles of LBT Data Accountability Legislation (LDAL)

We can utilize the thinking behind the FCRA data accountability approach to look towards legislation that holds LBT providers accountable for their use of the data they process and retain to provide their services. Such a proposed LDAL would actually likely be much simpler than the FCRA because of the different objectives and requirements of the two pieces of legislation. The FCRA is a complex piece of legislation, but the following characteristics, abstracted from [29], are relevant to a discussion about a proposed LDAL:

- **Limited, enumerated uses of data:** The FCRA limits use and disclosure of information for an explicitly enumerated list of purposes. Such a provision would also be necessary under proposed LDAL, though careful thought would need to be given to the list of purposes for use/retention of location information in order to strike a reasonable balance between a user's expectation of privacy and a LBT provider's interest in innovating. Such balance would be necessary for support from both stakeholders.
- **Access to personal data:** The FCRA contains provisions for consumers to obtain the contents of their credit report and request changes to inaccurate information, as inaccurate information may have negative consequences for consumers. The consequences of inaccurate information for LBTs are likely to be minor, given that the types of decisions based on whatever personally-identified location information is available are also likely to be convenience-related (except for perhaps E-911 and similar services). Given the less-consequential nature of LBTs (at least for now), user access to personally identifying data does not seem to be a necessary requirement, and such a requirement could unduly burden LBT providers if the personally-identifying information contains proprietary information. Instead, a useful compromise would be to allow users to request that LBT provdiers erase, anonymize, or aggregate any personally-identified location data such that it is no longer personally-identified.

Legislation based on the above principles, to hold LBT providers accountable for their use of the data they process and retain to provide their services, would significantly improve the privacy of LBTs. While not unduly burdening LBT providers or restricting their ability to innovate, such legislation would enable users to take advantage of the convenience of LBTs with the confidence that they were assuming reasonable levels of privacy risk.

5.2 Corporate Policy Approaches

We take "corporate policy approaches" to LBT privacy to mean measures that LBT providers, vendors, etc. should implement to protect user privacy and ensure that users of LBT products can make informed decisions about their privacy risks. Note that policies may ultimately manifest themselves in a number of ways, including notice to the user (the traditional "Privacy Policy"), internal procedures, technical measures, etc.

We begin by discussing an industry-wide policy recommendation, and then structure our own recommendations around two of the key tenets mentioned at top of this section: providing users with *information* and *choice*.

5.2.1 Industry-wide Policy Recommendations

While interviwing Jane Horvath, Senior Privacy Counsel, Google Inc., we learned about an industry-wide best-practices guide for providers implementing LBTs. Provided by the CTIA – The Wireless Association, the *Best Practices and Guidelines for Location-Based Services* [26] document provides guidelines for how LBT providers should deal with the privacy issues associated with LBTs. Interestingly, the CTIA guidelines focus on notice and choice — the same principles that we arrived at (arguably) independently before being shown the CTIA guidelines.

It is worth noting that compliance with these guidelines is indicated by "self-certifying" – that is, a provider who believes they have followed the guidelines appropriately may include the statement "`<LBS provider name>` follows CTIA's Best Practices and Guidelines for Location-Based Services"[26]. In our interview with Google, Jane Horvath mentioned that Google is now in compliance with the CTIA best practices, but they have yet to update their privacy policy/etc. with this notice. A cursory search for other "self-certifying" implementers by using Google to search for the string `"follows CTIA's Best Practices and Guidelines for Location-Based Services"` of the CTIA best practices turned up no results.

Our recommendations are more idealistic and comprehensive (a luxury we have), while CTIA guidelines appear to be focused on implementation amongst widely-varying industry providers. Overall however, our recommendations generally seem compatible with the CTIA best practices, and we would recomend implementation of these best practices as a minimum for any LBT provider. It will be interesting to see if this relatively new baseline gains traction as the LBT industry matures.

5.2.2 Informing Users

Informing users of their privacy risks when using any service that is not plainly understandable is difficult. There is a delicate balance that must be maintained in providing information for users to assess their privacy risks. On the one hand, providing users with too little or overly vague information to assess their privacy risks may cause users to be skeptical or uncertain about, or even unable to accurately assess their privacy risks. On the other hand, providing users with too much or or overly complex information to assess their privacy could lead to non-technical users, arguably the majority of users, becoming overwhelmed or paranoid. Imbalance to one extreme or the other have the potential to affect the market for a LBT application, device, or service.

The concern over managing privacy information is clear from practice too. Videos on Google's privacy[3] YouTube channel provide a clear "Don't Panic"[4] messsage to non-technical users, while more advanced users may find that the videos gloss over potentially important points. In communicating with Skyhook about an interview, we found that they would only communicate with us in writing, via email, to ensure there was no confusion about the message they were communicating with respect to privacy, given that they "[had] been subject to some spurious claims in the past."

In tension with the need for managing privacy concerns to ensure marketability of a LBT product, there is an ethical obligation on the part of a LBT provider to ensure that they have informed consent of the user [26]. The privacy policy or notice is typically the instrument for informing the user in order to facilitate obtaining informed consent. The policy notice must contain sufficient detail, and be written with sufficiently clarity, that an average user is able to comprehend the notice and understand how the concepts will affect their expectation of privacy.

5.2.2.1 Improving Privacy Policies

Unfortunately, many commonly-available LBTs do not provide sufficiently informative privacy policies to allow users to actually grant, in our opinion, truly informed consent. Based on our research conducted in preparing this paper, we offer the following recommendations, with examples, for improving LBT privacy policies to better inform users of their privacy risks when using LBTs.

[3] `http://www.youtube.com/googleprivacy`

[4] Hitchhiker's Guide to the Galaxy

Distinct from other notices To allow users to easily locate and read the privacy policy, it must be identified separately from any other "Legal" or "Terms of Use" notice for a LBT.

The Apple iPhone is a particularly bad example of this, burying the skeletal privacy policy in the middle of its legal terms of use[5] and GPL license notices.

Unique for a particular LBT To allow users to understand their privacy risks by using a particular LBT, the privacy policy must be written for *that technology*, instead of displaying or linking to a generic corporate privacy policy. This is particularly important for devices such as the Apple iPhone or T-Mobile G1 that utilize several LBTs in concert as part of their infrastructure.

Again, the Apple iPhone is a particularly bad example of this. The mention of Apple's privacy policy as referred to above is actually a hyperlink to Apple's generic *corporate* privacy policy[6].

Up-to-date and maintain a revision history To allow users to accurately determine their privacy risks at the time of using a LBT, the privacy policy must be up to date, and new technological changes should be implemented without also updating the privacy policy. To disclose to the user any changes that have occured to the privacy policy over time, the privacy must have a revision history that allows the user to view and compare the current privacy policy with previous versions of the policy.

With respect to the revision history, Google.com provides a good example of this[7]. Unfortunately, this functionality is only present on the full Google privacy policy, not on any of the particularly relevant product privacy policies such as Google Maps or Google Mobile.

Written in plain language To allow non-technical/non-lawyer users to understand the privacy risks of using a particular LBT, the privacy policy should be written clearly in plain language.

Google provides a good example of this in nearly all of its privacy policies, including its full privacy policy[8] and Google Mobile privacy policy[9]. Google also goes further by

[5] similar to `http://store.apple.com/Catalog/US/Images/iphone_tcs.pdf`

[6] `http://www.apple.com/legal/privacy/`

[7] `http://www.google.com/intl/en/privacy_archive.html`

[8] `http://www.google.com/intl/en/privacypolicy.html`

[9] `http://www.google.com/mobile/privacy.html`

providing a higher-level "Privacy Overview"[10] that is not specific to any particular product, but rather describes the general privacy behaviors of all Google products.

State clearly how a user's data is collected, used, and stored While perhaps difficult, especially when striving for plain language, the privacy policy should describe specifically what data is collected from a user, and how it is collected, used, and stored.

Skyhook Wireless provides a particularly good example of a clear and detailed privacy policy for its WPS WiFi geolocation service[11]. In contrast, while most of Google's privacy policies are very clear and easy to understand, they do not specify clearly what happens to a user's location information. Terms including "such as", "may" and "sometimes" are common in Google's privacy policies.

This brings up an interesting tension, in that it is extremely difficult to develop a privacy policy that is technically clear for advanced users, and yet readable and written in plan language. An ideal approach might be to provide two "levels" of privacy policy — one that is technically developed enough for advanced users to assess their risks to their satisfaction, and another that is simplified for non-technical users who are interested in a general understanding of the privacy implications of a particular LBT.

Explicity identify and link to other parties with whom data is shared The privacy policy must explicitly identify other parties with which it shares data and the nature of the data shared, as well as providing a link to the privacy policy of the third party. Given the multi-provider nature of LBTs where location data may be transferred between multiple LBT providers in a single transaction, this point is particularly important.

This point is also particularly important for LBT devices like the iPhone or G1 because these devices will typically be the originating point of these multiple provider interactions. Neither provider offers an excellent example of the type of explicit privacy policy we would like to see, but Google provides a somewhat reasonable example for the G1[12], where near the end of the policy it lists other Google products used by the G1 (such as Gmail) and provides (non-hyper)links to those product privacy policies.

While better than the iPhone in identifying the other LBTs it uses (see Chapter IV) the G1 privacy policy still does not provide a holistic, easy-to-understand privacy policy for the device and its applications — it forces users obtain an understanding of the "integrated"

[10] `http://www.google.com/intl/en/privacy_highlights.html`

[11] `http://skyhookwireless.com/whoweare/privacypolicy.php`

[12] `http://www.google.com/mobile/android/privacy.html`

privacy policy that exists when using multiple LBTs together. To better represent the overall picture of privacy for LBTs, and in particular LBT devices, we propose an alternate method of communicating privacy policies: a graphical privacy policy.

5.2.2.2 Graphical Privacy Policy

While looking at LBTs in greater depth, and in particular when studying their privacy policies, it became clear that it is very difficult, even for advanced users, to fully understand the privacy implications of using LBTs by reading their privacy policies. This was primarily due to the fact that the LBTs we focused on the most – the Apple iPhone and T-Mobile G1 – involve multiple LBT providers in single transactions.

There are various existing approaches to simplifying privacy policies for the comprehension of average users. As mentioned before, Google in particular emphasises the use of plain language in their privacy policies, and has also produced videos explaining the privacy policies and controls of their products in a video and multimedia presentation style.

In examining user responses to privacy policies, Milne [30] notes:

> ...[O]nline privacy notices are documents that can be presented in a hyper-text format. As a result, there are opportunities to promote better comprehension of online notices through the use of links than in the offline world. The open-ended comments and other evidence such as Hochhauser (2001) suggest that many current online privacy notices do not exploit these opportunities. If consumers perceive privacy notices as being irrelevant because of format issues, they may balk at even attempting to read them....

As we noted in the previous section, hyperlinks between the multiple providers utilized by a LBT would be helpful in allow the user to better comprehend the "integrated" privacy policy that exists when using multiple LBTs together. However, to take Milne's findings one step further, we propose an interactive, graphical means of displaying privacy policies to illustrate the sometimes-complex nature LBT privacy to the user.

A first example of this graphical privacy policy approach is shown in figure 5.1. This is an example of a diagram that illustrates all of the location information communication relationships on a particular device — in this case a hypothetical iPhone configuration with hypothetical second party LBT providers that fulfill the Core Location API and a "third party" application, Urbanspoon, installed on the device. While this diagram does not not illustrate the intimate details of each interaction with other parties, it provides an overview to give the user a general understanding of the communication relationships of their LBT device.

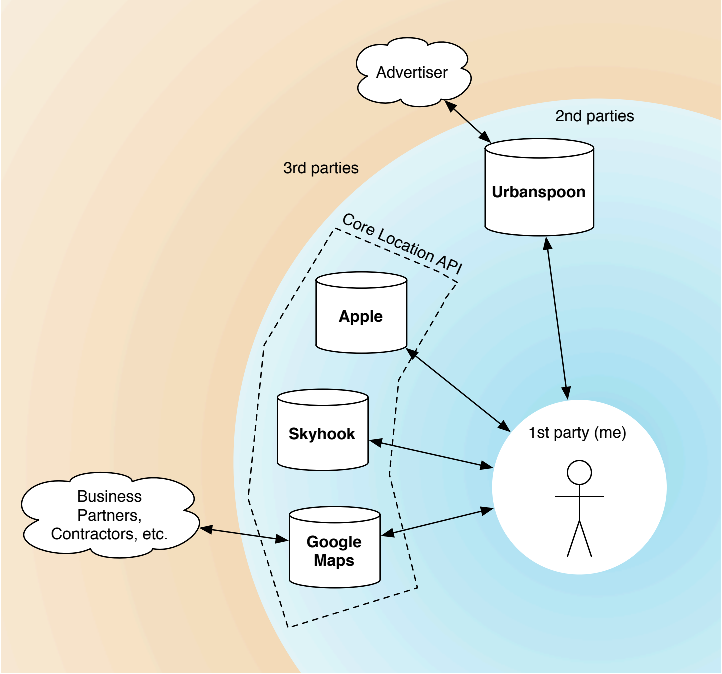


Figure 5.1: A simplified collective privacy diagram for a LBT device (in this case a hypothetical iPhone configuration). The diagram illustrates the number of LBT providers involved, and their interaction with other third party entities whose privacy policy is not included with/agreed to along with the rest of the LBT device

The interactive nature of this proposed presentation method now becomes helpful. By clicking on a particular LBT provider – in this case Google Maps – the user's graphical representation would change to that of figure 5.2.

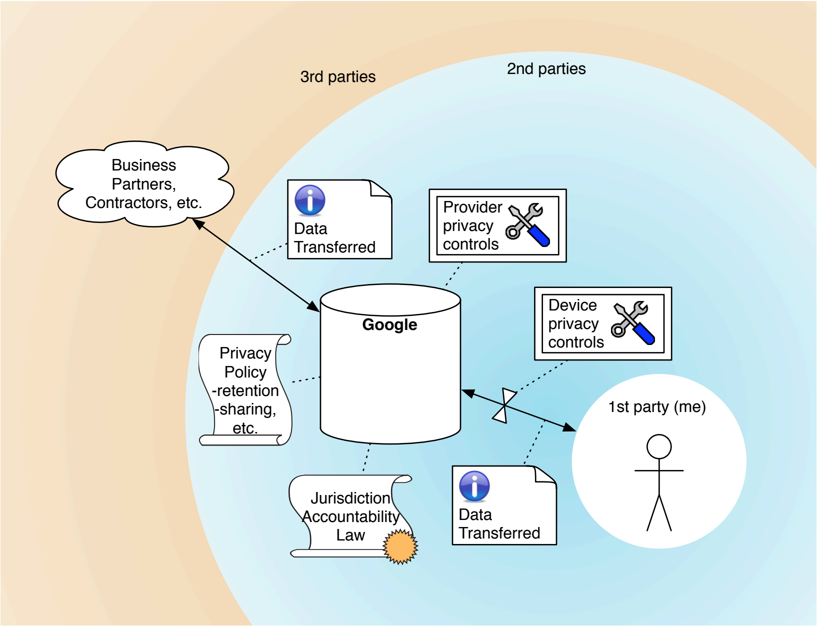

3rd parties
2nd parties
Business Partners, Contractors, etc.
Data Transferred
Provider privacy controls
Google
Device privacy controls
Privacy Policy -retention -sharing, etc.
1st party (me)
Jurisdiction Accountability Law
Data Transferred

Figure 5.2: A detailed privacy diagram focused on one (hypothetical) LBT provider but offering more information about that provider and its policies, as well as its relationship with third parties

This more detailed representation focuses on the specific details of a data/provider relationship between the user and one particular LBT provider, effectively providing a graphical "translation" of the textual privacy policy. The graphical model of the privacy policy represents and allows the user to determine

- What data is collected from the user by the LBT provider ("Data Transferred" in the figure above)
- What the LBT provider's privacy policies are with respect to data retention, sharing, aggregation, etc.
- What, if any, jurisdictional accountability law governs the LBT provider
- What, if any, data is transferred on beyond the LBT provider, and to whom it is transferred

The diagram also illustrates control points, to show the user how and where they can control their privacy risks through control of data transfer, storage, personal identification, and other features. Ideally, the interactive nature of this graphical representation would allow a user to control their privacy options directly through the image, or via hyperlink.

This graphical or diagrammatic approach to depicting a privacy policy appears to be an uncommon (or dare we say unique) approach in the literature and in practice. The only references found relating privacy policies to diagrams were for data flow diagrams used by software engineers in communicating about privacy risks and assessments of professional network infrastructures.

5.2.3 Offering users choice

In matters of user choice and control over their privacy risks, our thinking diverges from Google, CITA – The Wireless Association, and others. Most LBT providers, as described in privacy policies and elsewhere, believe that choice is important, but feel that it is sufficient to provide a binary choice — use a technology and increase privacy risks, or do not use the technology and avoid increaing privacy risks. We believe that this approach to considering choice is unfair and potentially unethical, and so finer-grained options must be provided to offer users *real choice* in determining their privacy risks.

As we've discussed previously, there is a tension between a user's desire to participate in technical innovation or enjoy convenient new technologies, and their willingness to increase their privacy risks in doing so. Due to asymmetric information about privacy risks, potentially resulting from a vague or difficult-to-comprehend privacy policy, potentially coupled with marketing of slick LBTs that promote their features/convenience/etc. without promoting their risks, users are placed in a position of unequal bargaining power with the LBT provider. In such a position, while users may technically provide consent to the provider to utilize their technology, we argue that consent is not fully informed, and thus the user was "unethically" forced to make a choice [31].

In this situation, users faced with a binary choice in participation are placed in a lose-lose situation — opting-out protects privacy at the cost of participating in convenient/fun/etc. technological innovation, while opting-in allows the user to enjoy participating in technological innovation at the cost of increased privacy risks associated with the use of this technology. To avoid this lose-lose situation, LBT providers should strive to provide fine-grained choice to users. What this means depends on the type of location information involved, but typically it would allow the user to disclose some but not all of the information about their location, or to disclose location only for certain purposes or only to certain providers/parties.

While offering fine-grained choice to users may pose significant technical challenges for LBT providers, it would also likely benefit providers in the long run because of increased potential for adoption and credibility if/when user concerns about the privacy of LBTs overtake the perceived benefits of these technologies. An anecdotal example of this is the evolution of Facebook's privacy controls from its early growth in 2005 to present. Initially minimal, binary privacy controls were eventually replaced by the current, extremely fine-grained approach now used by Facebook, where users may control which users/groups of users/networks/etc. can view elements of their profile if they so desire.

5.3 Technical Approaches

While the legal and corporate policy approaches to minimizing user risk when interacting with location-based services are likely the most powerful measures, technical measures can also significantly contribute. Here, we outline two categories of technical measures: changes to the mobile device platform itself and changes to the infrastructure behind location determination services.

5.3.1 Changes for Devices

In order for the user to be fully informed about the sprawling network of privacy policies on the device, every privacy policy should be readily available on the device for the user to explore. If the privacy policy for the operating system needs to refer to a privacy policy for another service, it should be hyperlinked so that the user only needs to click or tap on the link to view the other policy. Another source for user confusion is from the myriad of third party applications. Each application can potentially have a different privacy policy governing user data. A central location in the device's settings menu could easily be set aside to enable the storage of application privacy policies. When an application is installed, it can also place an entry in the application policy section, allowing the user to easily browse the individual policies for each application on his device.

Also, control over the individual access to location data by each application needs to be more fine-grained. Currently, most devices only allow a per-application decision at install time or when the application is first run, and they only allow users to disable location determination globally instead of for individual applications. First, users should be able to view a list of all applications and individually determine (at any time) the access each application has to location information, allowing easier user choice. Second, each application does not necessarily need to know the *exact* coordinates of the user; some need a neighborhood-level location, and some only need a city-level location. In the same setting panel where users can determine individual application access to location information, users should also be able to choose the level of detail each application can access, allowing some application less fine-grained access to location data.

5.3.2 Changes for Location-service Providers

Regardless of any external access to the location database of a location determination service, it is possible for a malicious user to masquerade as a specific access point and allow a mobile device to only see that access point. Then, the location determination service will return the best guess for the device's location, which will naturally be the estimated location of the faked access point. In this way, an attacker can determine the location of any access point. If the service providers, however, required that a mobile device listed enough access points in the same geographic area to be confident that device really was in that area, the effectiveness of this attack could be minimized. For example, if there are five access points in the database that are tightly clustered together, and a mobile device only reports one of those access points, the service can reasonably conclude that the device is not actually in that location and refuse to give a location estimate. This works well in urban environments where the probability of legitimately seeing only one access point is low, but does break down in rural areas where access points may be very spread apart. In these cases, the system would need to allow a location determination based on only one access point, which is a "necessary evil" in order for the system to operate.

Furthermore, any service which provides or utilizes location should encrypt the network connection where a location is transmitted. Using SSL does increase the computational overhead on the server side of the connection, but not significantly enough to justify revealing a user's location over the network.

Also, any service utilizing location should take a cue from Skyhook's privacy policy and ensure that user data is fully anonymized before it is stored for any purposes. Even if location data is stored with a random identifier that cannot be tied to a particular user, but the identifier persists across multiple location queries, it is easy to reidentify the user based on some location data (inference of home and work locations, for example) and then use that identification to track a user. Location data only needs to be logged for purposes of training a positioning database, and this does not require a unique identifier for each entry. Thus, unless a provider is explicitly providing location-logging services to the user, an identifier of any kind should not be stored on a provider's server for any length of time beyond what is need to finish a particular transaction.

Conclusion

Next-generation location-based technologies pose a number of privacy risks to the users of these services. Even more importantly, the bulk of popular location-based technologies currently available today pay little to no attention to the need for users to be able to accurately assess the privacy risks that they expose themselves to in using these technologies, nor do they offer users many options in controlling these risks. While all providers *do* have privacy policies that concerned users can be directed to, most of these policies do not provide any reasonable level of detail for even technical users to assess the risks of using these location-based technologies.

While many of these technologies, and indeed the whole location-based technology market, is relatively nascent and still developing, this lack of concern for providing reasonable information and controls for privacy risks is troubling. Perhaps providers assume that users don't care that much about the privacy of their current location, instead being more interested in how many songs they can fit on their new smartphone, or how easy it is to find a nearby coffee shop. However, such a cavalier attitude on the part of providers and users is dangerous — ignorance will not maintain a user's expectation to privacy. This attitude is especially concerning given the explosive 3000% growth expected in the location-based technologies market over the next five years [3]. If privacy is not a concern for most providers while products are young and trying to gain footholds in the market, why should we expect it to become a concern and be dealt with later when providers and products are entrenched?

In light of the privacy issues illustrated in this paper, we propose solutions to both the direct privacy risks from the use of location-based technologies, as well as for the larger issue of provision of information and options for user choice in managing their privacy risks. We propose legal measures that will hold user location data to a higher standard of judicial scrutiny, and that would utilize use-based controls over the large amount of information collection inherent to modern location-based technologies. We propose policy measures that providers should implement to ensure adequate information is disclosed to users to allow risk assessment, as well as properly granular choices to allow users to manage their privacy risks without simply chosing to not use location-based technologies and thus not enjoy the benefits they provide. Finally, we propose technical measures to facilitate user choice in managing their privacy risks, as well as to prevent abuse of certain types of LBT provider infrastructure that can be abused to obtain potentially-identifying information.

Appendix A

Correspondence

The following questions were sent to or asked of the LBS providers indicated.

A.1 Google

After making contact with Jane Horvath, Senior Privacy Counsel for Google in mid-November, we asked the following questions during an interview on 9 December 2008.

1. General:

   - How does Google define "privacy"?
   - What is Google's general approach or philosophy to privacy? What motivates this approach? (i.e. law, market, customers, "don't be evil") What are your major priorities with regard to privacy?
   - What are your thoughts on the current state of privacy protection and information in the industry, both in general and especially with respect to location-based technologies? Where does Google stand, and what (if any particular) role does it play in the industry with respect to privacy?
   - Why is location privacy "high on your priority list"?
   - What do you see as questions/issues related to location privacy in the future? Do you have any thoughts on changes/policies/legislation for whole industry with respect to privacy, both generally and with respect to location privacy?

2. Understanding and controlling privacy risks:

- How do you deal with communicating potentially complicated privacy notices to users? Are you concerned with user comprehension of privacy policies?
- What do you think about granularity in user control of their privacy risks? Is it “fair” to only give the user the option to opt-out of using a service?

3. Geolocation:

   - Google seems fairly quiet about their Geolocation service used by (we believe) Android and other services[1]. How can users find the privacy policy for this service?

4. Mobile Search:

   - Does Google retain search queries associated with geographic locations? Are they associated with unique identifiers?
   - How would Google respond to a law enforcement request for stored search queries from a particular phone? (A particular phone being identified by IP, MAC, or some other unique ID.)
   - How would Google respond to a law enforcement request to filter search queries as they come in and log the queries from a particular phone?

[1] `http://google-code-updates.blogspot.com/2008/08/two-new-ways-to-location-enable-your.html`, `http://google-code-updates.blogspot.com/2008/10/introducing-gears-geolocation-api-for.html`

A.2 Apple

Not optimistic about receiving a response from Apple's press contacts, we sent the following mail to the contact address noted in Apple's privacy policy[2] on 3 December 2008. Unfortunately, we did not receive a reply before deadline. For the record, Stephen *does* own an iPhone.

```
From: "Stephen Woodrow"
To: privacy@apple.com
Subject: iPhone Privacy Questions

Hello,

I am an iPhone owner and recently have had some questions about the
privacy of information about my location when using my iPhone:

First, I tried to find an iPhone privacy policy to answer my
questions, but I could only find references to the general Apple
privacy policy (http://www.apple.com/legal/privacy/), which wasn't
detailed enough for my interests. Is there a more specific privacy
policy for the iPhone where I can find the answer to my questions?

Second, I was wondering what happens when i push the "Locate me"
button in the iPhone Maps application? How is my location determined,
and where is information about my location sent? A computer science
friend of mine suggested that Apple is using Skyhook Wireless to
determine location when using the iPhone. Does that mean I can refer
to Skyhook's privacy policy?

Third, I also have a couple of iPhone applications that I was
considering using (Urbanspoon is one example), but I was wondering how
I can determine what will happen to my location information when I use
applications I download from the iTunes store? Do these applications
have privacy policies?

Finally, I was wondering what I can do to control the use of
information about my location? I know about the dialog box that I can
click "don't allow" access to my location information, but are there
```

[2] http://www.apple.com/legal/privacy/

any other ways to control what happens with my location?

Thank you,
Stephen Woodrow

Bibliography

[1] C. R. Yinger, "Operation and Application of the Global Positioning System," *Crosslink: The Aerospace Corporation magazine of advances in aerospace technology*, vol. 3, no. 2, Summer 2002. [Online]. Available: http://www.aero.org/publications/crosslink/summer2002/02.html

[2] GPS III. Last accessed: November 23, 2008. [Online]. Available: https://buy.garmin.com/shop/shop.do?pID=72

[3] A. Sharma and J. E. Vascellaro, "Companies Eye Location-Services Market," *The Wall Street Journal*, November 21 2008, last accessed: 22 November, 2008. [Online]. Available: http://online.wsj.com/article/SB122722971742046469.html

[4] *Illinois v. Gates*, 1983. [Online]. Available: http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=us\&vol=462\&invol=213

[5] S. R. Strom, "Charting a Course Toward Global Navigation," *Crosslink: The Aerospace Corporation magazine of advances in aerospace technology*, vol. 3, no. 2, Summer 2002. [Online]. Available: http://www.aero.org/publications/crosslink/summer2002/01.html

[6] P. Massatt and W. Brady, "Optimizing Performance through Constellation Management," *Crosslink: The Aerospace Corporation magazine of advances in aerospace technology*, vol. 3, no. 2, Summer 2002. [Online]. Available: http://www.aero.org/publications/crosslink/summer2002/03.html

[7] C. Renso, S. Puntoni, E. Frentzos, A. Mazzoni, B. Moelans, N. Pelekis, and F. Pini, *Mobility, Data Mining and Privacy.* Springer Berlin Heidelberg, 2008, ch. Wireless Network Data Sources: Tracking and Synthesizing Trajectories, pp. 73–100. [Online]. Available: http://www.springerlink.com/content/g4502508jpp4xn04

[8] Garmin Ltd. – Company History. [Online]. Available: http://www.fundinguniverse.com/company-histories/Garmin-Ltd-Company-History.html

[9] I. Google. Google Announces Launch of Google Maps for Mobile With "My Location" Technology. Press Release. [Online]. Available: http://www.google.com/intl/en/press/annc/20071128_maps_mobile_my_location.html

[10] Z. Ji and R. Jain. (2008, June 6) Google enables Location-aware Applications for 3rd Party Developers. Google Mobile Blog. [Online]. Available: http://googlemobile.blogspot.com/2008/06/google-enables-location-aware.html

[11] K. Jones and L. Liu, "What Where Wi: An Analysis of Millions of Wi-Fi Access Points," *Portable Information Devices, 2007. PORTABLE07. IEEE International Conference on*, pp. 1–4, May 2007.

[12] I. Skyhook Wireless. Skyhook Wireless: How It Works: Wi-Fi Positioning. Last accessed: November 20, 2008. [Online]. Available: http://www.skyhookwireless.com/howitworks/wps.php

[13] C. Wiles. (2008, October 21) Introducing the Gears Geolocation API for all laptop WiFi users. Google Code Blog. Last accessed: November 23, 2008. [Online]. Available: http://google-code-updates.blogspot.com/2008/10/introducing-gears-geolocation-api-for.html

[14] Apple Investor Relations: Earnings Releases. [Online]. Available: http://www.apple.com/investor/

[15] Skyhook in Action: Location Apps. [Online]. Available: http://skyhookwireless.com/inaction/locationapps.php

[16] C. Foresman, "Apple rejects another app for 'duplicating functionally'," *Ars Technica*, September 22 2008. [Online]. Available: http://arstechnica.com/journals/apple.ars/2008/09/22/apple-rejects-another-app-for-duplicating-functionality

[17] iPhone SDK Agreement. Apple Inc. [Online]. Available: http://blog.wired.com/gadgets/files/iphone-sdk-agreement.pdf

[18] C. Moor, "T-Mobile G1 Event Round-up," *Talk Android*, September 23, 2008. [Online]. Available: http://www.talkandroid.com/260-t-mobile-g1-details/

[19] (2007, November 7,) Industry Leaders Announce Open Platfor for Mobile Devices. Open Handset Alliance. [Online]. Available: http://www.openhandsetalliance.com/press_110507.html

[20] Android Developer Challenge Gallery. Google. [Online]. Available: http://code.google.com/android/adc_gallery/index.html

[21] Android: android.location. Open Handset Alliance. [Online]. Available: http://code.google.com/android/reference/android/location/package-summary.html

[22] (2008, November) G1 Privacy Policy. Google. Last Accessed: December 9, 2008. [Online]. Available: http://www.google.com/mobile/android/privacy.html

[23] Skyhook Wireless: Who We Are. Skyhook Wireless. Last accessed: December 9, 2008. [Online]. Available: http://skyhookwireless.com/whoweare/

[24] Steve. (2008, September 10,) Get the physical location of wireless router from its MAC address (BSSID). Last accessed: December 9, 2008. [Online]. Available: http://coderrr.wordpress.com/2008/09/10/get-the-physical-location-of-wireless-router-from-its-mac-address-bssid/

[25] (2008, January) Skyhook Wireless, Inc. Privacy Policy. Skyhook Wireless. Last accessed: December 7, 2008. [Online]. Available: http://skyhookwireless.com/howitworks/privacypolicy.php

[26] CTIA. (2008, 2 April) Best Practices and Guidelines for Location Based Services. CTIA – The Wireless Association. [Online]. Available: http://files.ctia.org/pdf/CTIA_LBS_BestPracticesandGuidelines_04_08.pdf

[27] J. Horvath. (2008, July 9) Testimony of Jane Horvath, Senior Privacy Council, Google Inc., at hearing on the “Privacy Implications of Online Advertising”. U.S. Senate Committee on Commerce, Science & Transportation. Last accessed: December 8, 2008. [Online]. Available: http://commerce.senate.gov/public/_files/JaneHorvathGoogleOnlinePrivacyTestimony.pdf

[28] D. J. Weitzner, H. Abelson, T. Berners-Lee, J. Feigenbaum, J. Hendler, and G. J. Sussman, “Information Accountability,” *Communications of the ACM*, vol. 51, no. 6, June 2008.

[29] Electronic Privacy Information Center (EPIC). The Fair Credit Reporting Act (FCRA) and the Privacy of Your Credit Report. Last accessed: 9 December 2008. [Online]. Available: http://epic.org/privacy/fcra/

[30] G. R. Milne and M. J. Culnan, "Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices," *Journal of Interactive Marketing*, vol. 18, no. 3, pp. 15–29, July 2004. [Online]. Available: http://doi.wiley.com/10.1002/dir.20009

[31] A. Valys. (2007, 12 December) Privacy at Speed. [Online]. Available: http://groups.csail.mit.edu/mac/classes/6.805/student-papers/fall07-papers/privacy-at-speed.pdf