A Privacy Threat Model for Seattle Residents

By Adam Shostack, with substantial help from many Seattle Privacy Coalition members.

Summer/Fall 2017

TODO:

* Visually align tables

# Executive Summary

In May, 2017, we announced a project, “A Privacy Threat Model for The People of Seattle,” designed to create:

*“...a threat model of a day in the life of a resident in Seattle, and the way data is collected and used; that is, the potential threats to their privacy. A typical threat-modeling approach would focus on a system that we’re building, analyzing, or testing. In this model, we focus on the people, the ‘data subjects.’”*

Since then, we built the threat model. It produced interesting and unexpected results. For example, there is more privacy in using a car service, such as Lyft or Uber, than in driving your own car. (Your car is tied to you, and has license plates and potentially many radios including toll-payment, tire-pressure monitoring, and services such as On-Star. For details, see “How We Learned It.”)

We’re sharing this document because it’s reached a useful milestone and enables additional work by defenders, by researchers or hackers, and by policymakers. This threat model is a resource for anyone writing or assessing a “self-defense guide.” If you’re going to make investments in protecting your privacy, you need to ensure your defenses don’t leave easily exploited gaps. This threat model can be read as a catalog of problems that need solving by researchers, hackers, or entrepreneurs. It can be read as a catalog of privacy problems for policymakers. (A parallel effort to assess the privacy threats to a resident of a city in Germany or somewhere else with strong privacy law might lead to fascinating comparisons.)

This paper consolidates details from a set of blog posts about the project, explains the data we produced, and discusses potential next steps.

Table of Contents

[Executive Summary 1](#_Toc502914035)

[What We Learned: Summary 2](#_Toc502914036)

[How We Learned It 3](#_Toc502914037)

[Threat Modeling: a brief overview 3](#_Toc502914038)

[Threat Modeling for Seattleites 3](#_Toc502914039)

[Privacy Tradeoffs In a Commute 5](#_Toc502914040)

[The Raw Data 7](#_Toc502914041)

[Black Hole Sun (Alpha release) 8](#_Toc502914042)

[Conclusions and Next Steps 9](#_Toc502914043)

[Next Steps 9](#_Toc502914044)

[Glossary 11](#_Toc502914045)

[About Seattle Privacy Coalition 12](#_Toc502914046)

[References 13](#_Toc502914047)

# What We Learned: Summary

* We learned about threats we hadn’t previously considered. This was true for every participant, which is surprising in and of itself: there are few better-educated sets of people than those willing to commit hours of their weekends to threat-modeling privacy.
* We built a new way to contextualize the decisions we might make, evidence that we can generate these in a reasonable amount of time, and an example of that form.
* We learned about how long contextualizing would take (a few hours to generate a good list of threats, a few hours per category to understand defenses and tradeoffs), and how to accelerate that. (We spent a while diving deeply into threat scenarios in a way that didn’t help with the all-up models.)
* We saw how deeply and complexly mobile phones and apps play into privacy.
* We got to some surprising results about privacy in a typical Seattle commute.

Please note that company and product names are intended as evocative examples, not to comment on the privacy practices of those products or organizations.

# How We Learned It

This section presents an overview of typical approaches to threat modeling, distinguishes them from this project, explains what happened and what we learned (both about the privacy threats, and about this approach to privacy threat modeling), and then presents our results.

## Threat Modeling: a brief overview

Typically, a modern threat modeling process focuses on a software product or service that is being developed. It can focus on the entirety of the product[[1]](#footnote-1), or on new features for the next release. Threat modeling is usually carried out by the software engineers working on the product, and involves four key questions:

1. What are we working on?
2. What can go wrong?
3. What are we going to do about it?
4. Did we do a good job?

A product team trying to threat model a person’s day would find a lot of issues that are outside of their control, and so spend energy enumerating issues that they cannot fix. Producing this list of unfixable problems would sap energy from that product team. However, that doesn’t mean that a threat model cannot help us look at those things; we just need to carefully select our goals and participants.

Because threat modeling usually focuses on a specific product, it’s difficult to use the traditional process to assess the overall state of privacy, so we created a new process to try to do so.

## Threat Modeling for Seattleites

We designed and ran a new type of threat modeling process, shown in Box 1. We got a room in the local public library, and advertised the effort to our organization’s discussion list, and to the public, using social media. We met, introduced the idea, introduced ourselves to each other, and set some ground rules about respect, collaboration, and assuming good intentions. Next, we started capturing ideas on a whiteboard about “what happens in a typical day?”

BOX 1: Planned Approach

What is a Seattlite’s day?

What data is collected, how, and by whom?

What models can we create to help us understand? Is there a good balance between specificity and generality?

What can go wrong?

What could we do about it?

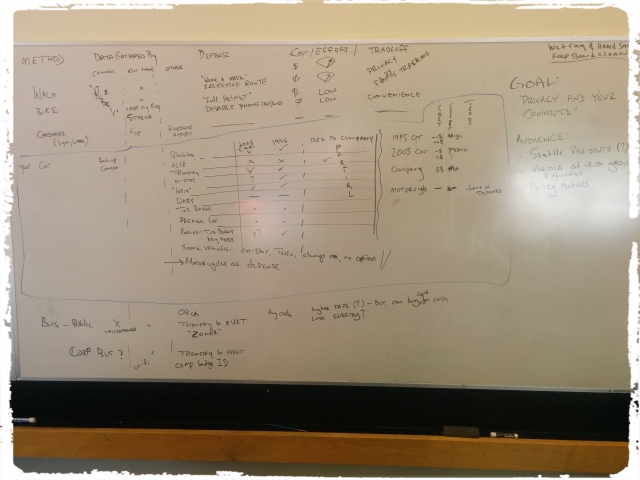
Did we do a good job?

Over the course of two hours, we discovered that answering the “who is collecting the data” question gets tricky – many details might play in. We ended up with “government” at various levels of specificity, and private companies with whom the Seattleite might or might not have a relationship. Often where the data goes was not obvious, or we had enthusiastic disagreement about how to characterize it. Over time, we moved to simply noting that the data was being collected.

We also got into a vigorous discussion of “is this mandatory?” There are levels of mandatory, ranging over:

* Completely optional (an internet-connected bathroom scale)
* Social pressure (share your running monitor with your running club)
* Work pressure (“wellness programs”)
* Financially incented (“Get a discount to do this”)
* Financially mandated (“Send data to god-knows where so your insurance company keeps paying for the CPAP machine that helps you manage your sleep apnea.”)
* Work required
* Government required

Figure 1: What can go wrong, on the whiteboard



We did not make extensive use of this data. Perhaps a different model would be more usable, and perhaps a different methodology would make better use of it.

We also had to pull out of very enthusiastic denunciations of “why would anyone do that?!”

The question of “what can go wrong” also turns out to be tremendously complex to answer. First, it depends on “for whom.” The same data collected about a political activist and a non-active person carries very different implications. People’s concerns change with gender, race, immigration status, religious beliefs, sexual orientations, age, or physical independence/need for assistance.

BOX 2: Executed Approach

What is a Seattlite’s day?

What data is collected and how?

What could they do about it?

In our second meeting, we ran into an issue: we had not defined an output state. Traditional threat modeling processes might result in a list of issues to be fixed, enumerated in a report or a set of software “bugs”[[2]](#footnote-2) to be addressed. This was intentional on my part: I wanted to see where things led us. It was also a mistake. Not having an output state led us in circles.[[3]](#footnote-3) We decided to focus on a list of ways data could be collected and a list that allows us to look at tradeoffs between various choices.

## Privacy Tradeoffs In a Commute

Our third meeting looked at the privacy tradeoffs between ways to commute, as shown in Figure 1 and transcribed in Table 1.

Table 1 has four columns: “Commute type”, “data gathered by”, “defense”, and “cost/effort/tradeoff.” To explain:

* *Commute type* is how someone regularly gets to work. Note that it’s a model, and we group things that are similar (Lyft/Uber/Taxi) even though they’re not identical. Selecting a good degree of generalization helps collapse the data into manageable chunks. Too little generalization results in many similar lines, too much and you lose actionability. Protecting your privacy requires consideration of each data-gathering method.
* *Data gathered by* is a mechanism, not the party who gathers the data.
* *Defense* is how you might defend yourself.
* *Cost/effort/tradeoff* is an assessment of what you might be spending, in cash, time, frustration, or other tradeoffs to use the defense.

|  |  |  |  |
| --- | --- | --- | --- |
| Commute type | Data Gathered by | Defense | Cost/Effort/Tradeoff |
| Bicycle | Cameras, Microphones (Public and/or private, e.g. coffeeshop, etc.) | ‘Wear a mask’ | Low/Medium/Facial or Voice ID |
| Your bike (not a bike share) |  | Selective Route | Zero/Medium-High/May lose fitness tracking or similar |
|  |  | Full Helmet | Medium/Low/Facial ID |
|  | Personal Cell Phone (Location, Apps such as MapMyRide, Strava) | Disable personal phone/Bluetooth/Wifi | Zero/Low/Convenience, ability to be in constant touch with relatives, etc. |
|  |  |  |  |
| Carshare (Lyft, Uber, Taxi) | Cameras, Microphones (if exists, in taxi or rideshare) | ‘Wear a mask’ | Unrealistic, likely can’t use almost any of these defenses save turning off own cell |
|  | Cell Phone (Location, route apps such as Waze) | Disable personal phone/Bluetooth/Wifi |  |
|  | Rideshare, taxi usage history | Disable vehicle monitors |  |
|  | Personal Cell phone/GPS/Blackbox/OnStar (or similar) of vehicle |  | NOTE: Some blackbox/OnStar are hard to disable, but we don’t know that they tie strongly to you |
|  |  |  |  |
| Public Bus/ Subway/Transit | Cameras, Microphones (Public in transit vehicle.) | N/A | N/A |
|  | ORCA card (or similar type of ‘smart pass’) | Pay Cash | Higher fare, potentially lose subsidy/ |
|  | Personal Cell Phone (Location, Apps) | Disable personal phone/Bluetooth/Wifi | Zero/Low/Convenience, ability to be in constant touch with relatives, etc. |
|  | Telemetry to fleet (e.g. Zonar, etc.) | N/A | N/A |
|  |  |  |  |

**Table 1 overall tradeoffs.**

What we can learn from this:

* Cell phones and apps are a “horizontal threat” and apply in each commute modality.
* Walking and biking are the most privacy preserving commutes. Everything else generates long-term records of your movement. However, some electric bikes have anti-theft GPS built in, as do the new dockless rental bikes.
* It’s easier to prevent camera tracking on a bike because a helmet is not as attention-grabbing as a mask. Bikes also limit “gait biometrics.”
* Motorcycles have fewer radios than a car, and are generally much simpler, but still carry license plates and may be tracked via road toll systems. There are obviously complex tradeoffs involved in motorcycle commuting, but it wasn’t obvious to us, going into this project, that privacy plays in those tradeoffs.
* Between Lyft/Uber and your own car, your own car is trackable in more ways, and more ways that tie to you.  Unless you’re worried about those companies, you’re better off with a taxi or carshare.  If you’re worried about the federal or local government, there are a lot of parties a government will subpoena, and so that’s neutral.  (Taxis vs app-driven: if you call a taxi, your pickup location/phone combo may be recorded.  If you hail it, then pay with a card, your drop-off location may be recorded.  If you hail and pay cash, then you’re more private than with an app.)

We also looked at phones. There’s a set of radios in a phone, some of which (Bluetooth, wifi) can be turned off with less impact on usability. The cellular network radios can only be turned off with a substantial loss of functionality. We also discussed differences in usability of turning off app access to location between various brands.

|  |  |  |  |
| --- | --- | --- | --- |
| Data Gathered by | Phone Type | Defense | Cost/Effort/Tradeoff |
| Ultrasound | Modern smartphone | No microphone access | Hard to turn off app-by-app |
| Bluetooth |  | Turn off Bluetooth | Toggles at OS level\* |
| Wifi |  | Turn off wifi | Toggles at OS level\* |
| Apps |  | Turn off per app | May differ by OS |
| Location | Android | Turn off location | Android users in TM session reported this is tough to keep set |
| Location | iOS | Turn off location | Easier on iOS |
| CSLI | Any | Airplane mode | High tradeoff |
| Stingray | Any | Airplane mode | High tradeoff |
|  | Android | Detection apps |  |

TABLE 2: Phone privacy tradeoffs

\* Note that in Apple IoS 11, the easily accessible toggles have surprising behavior. [EFF, Apple]

## The Raw Data

The raw data from this project is available under a creative-commons attribution license. The raw data is distributed in two forms: Microsoft Excel xlsx[[4]](#footnote-4) and HTM. Here’s a version [in Excel](https://adam.shostack.org/seattleprivacy/threatmodelforseattle_v1/Seattle%20Resident%20Threat%20Model.xlsx) and a web view, [exported HTM](https://adam.shostack.org/seattleprivacy/threatmodelforseattle_v1/Seattle%20Resident%20Threat%20Model.htm) here.

The spreadsheets also contain models of “your own private car”. Note there’s a bit of extra modeling in the “commute type” column about the type of data collection; we wanted to maintain similar headings.

# Black Hole Sun (Alpha release)

We no longer release programming languages and say “here’s my approach to programming,” because that would be confusing. It worked in 1960 to name something “A Programming Language,” and now we have many. Similarly, we have many threat modeling approaches.

The methodology covered here is “Black Hole Sun (Alpha).”

Black Hole Sun Privacy Threat Modeling consists of:

|  |  |  |
| --- | --- | --- |
| Stage | Task | Output |
| What are we working on? | Brainstorm activities | Activity list |
|  | Use activity list to create a model | Activity category list |
| What can “go wrong”? | Analyze each category of activity for data gathering | Who/what/how collects data |
| What can be done? | Tradeoff analysis | Tradeoff list |

# Conclusions and Next Steps

We have created, executed on, and reported on a new approach to privacy threat modeling, Black Hole Sun. We are sharing the full results because both the approach and the information are promising and worth pursuing.

## Next Steps

### Further Analysis of Initial Data

We have performed tradeoff analysis for the commute data we gathered. We or someone else could perform analysis of the remainder of the data.

### Generate Data for Other Locations

We analyzed the situation in Seattle. Other locations may have other data gathering. For example, the Perpetual Line Up project documents that 16 states share their motor vehicle licensing data with the FBI for warrantless facial recognition. [Perpetual Lineup] A New Jersey security researcher discovered that his EZPass was being read at various non-toll locations around New York. [Puking Monkey]

### Vulnerable Populations

The Black Hole Sun approach could be useful for identifying concerns and threats for vulnerable populations. For example, a targeted threat model could come out of a facilitated discussion among survivors of domestic abuse who have experience with coercive technologies.

The information we’ve gathered can support and accelerate analysis for other vulnerable groups. For example, we know that cell site location information can only be disabled by discarding a mobile phone or leaving it in airplane mode. We also know that DHS collects mobile phone information from DACA applicants. We have not attempted to analyze this or its implications, but we’d be happy to do so in partnership with organizations that have specific concerns.

### Technologists

Our analysis presents a large set of issues where the tradeoffs available are not satisfying. Research into these areas could produce new technology that creates new options.

For example, we have identified a set of privacy implications to modern cars. We hope manufacturers pick up on these problems and work to address them. We also know that there is a rich market of aftermarket improvements for cars: we hope that researchers and hackers start working on automotive privacy. We believe that this approach would yield smart results applied to any other object that might be “connected” and part of the “Internet of Things.”

### Policy Makers

The world is changing rapidly. Tracking and gathering data has quietly gone from an “Orwellian nightmare” to an unremarkable aspect of our lives. Much of this has been enabled by well-intended new systems deployed by governments or new requirements imposed by governments. There is an opportunity to work back from the data gathering systems we have identified and gain an increased understanding of the impacts of policy. The data we have produced also enables better analysis of new systems and requirements, and we hope that it is used for that.

# Glossary

***The Connector*** is Microsoft’s corporate bus system for employees. It uses a HID-type card to control who can ride, and ridership is tracked to control who can get a reservation. We use it because the Connector busses are visibly branded, we do not mean to imply anything about Microsoft’s privacy practices.

***Dockless bikes*** are rental bicycles made available in Seattle by several companies such as LimeBike or Spin. Bikes are rented via an app, and each bike is tracked via GPS. [Washington Post]

***Good To Go*** is the Washington State Department of Transportation’s toll transponder system, used to pay for several bridges and other toll roads, and operated by a variety of private companies. [Seattle Times] Those bridges and roads also photograph each license plate that transits [Mocek].

***ORCA Card*,** the King County Metro system’s ride card. ORCA cards are contactless smart cards used on entering and exiting the system. We did not consider issues of RFID attacks.

# About Seattle Privacy Coalition

About

The fundamental right of privacy for all human beings underpins free society, yet today this right is challenged by technological advances that open vast new opportunities for the aggregation and analysis of personal data.

We believe that by starting at the municipal level, we as individuals can have an impact and help re-establish meaningful privacy expectations that can inform and inspire the work of others. Seattle Privacy formed in March 2013 and incorporated as a non-profit corporation in the state of Washington in 2014.

Vision

Worldwide legal protections of privacy for all people, wherever they live, whatever their citizenship, against increasingly totalitarian government surveillance programs and intrusive and cynical corporate data collection.

Mission

To urge and empower the City of Seattle and other local governing bodies to take advantage of our region’s leadership in technology and progressive politics to lead the United States to restore and protect all people’s right to privacy.

Copyright: © 2017 Seattle Privacy Coalition. This document is provided under the Creative Commons Attribution 3.0 United States License. ([CC BY 3.0 US](https://creativecommons.org/licenses/by/3.0/us/)).

# References

[Apple] “Use Bluetooth and Wi-Fi in Control Center with iOS 11” Apple, Oct 23, 2017. https://support.apple.com/en-us/HT208086

[EFF] “iOS 11’s Misleading “Off-ish” Setting for Bluetooth and Wi-Fi is Bad for User Security,” Andrés Arrieta, EFF blog, October 4, 2017. <https://www.eff.org/deeplinks/2017/10/ios-11s-misleading-ish-setting-bluetooth-and-wi-fi-bad-user-security>

[Mocek] “WSDOT answer questions about privacy and records of SR 520 bridge crossings,” Phil Mocek, May 11, 2011 https://mocek.org/blog/2011/05/11/wsdot-re-records-of-sr520-bridge-crossings/

[Perpetual Line Up] “The Perpetual Line-Up”, Clare Garvie, Alvaro Bedoya, Jonathan Frankle, Center on Privacy & Technology, Georgetown Law, <https://www.perpetuallineup.org/>

[Puking Monkey] “The Road Less Surreptitiously Traveled”, Puking Monkey, Defcon 21, and “‘Intrusive and Unsettling’: Man Rigs E-Z Pass Toll Tag to Show It Tracks Cars Outside of Tolls,” Liz Klimas, The Blaze, September 20, 2013. <http://www.theblaze.com/news/2013/09/20/intrusive-and-unsettling-man-rigs-e-z-pass-toll-tag-to-show-it-tracks-cars-outside-of-tolls>

[SeattleTimes] “Computer glitch causes hefty late bills for WSDOT tolls,” Bob Young · April 21, 2016, <https://www.seattletimes.com/seattle-news/transportation/computer-glitch-causes-hefty-late-bills-for-wsdot-tolls/>

[Washington Post] “Will dockless bike services boost bike commuting? Cities are betting on it,” Luz Lazo, December 2, 201 https://www.washingtonpost.com/local/trafficandcommuting/will-dockless-bike-services-boost-bike-commuting-cities-are-betting-on-it/2017/12/02/4fd6216c-c895-11e7-b0cf-7689a9f2d84e\_story.html

Original blog posts: “[A Privacy Threat Model for the People of Seattle](https://adam.shostack.org/blog/2017/05/a-privacy-threat-model-for-the-people-of-seattle/),” Adam Shostack, May 8, 2017, https://adam.shostack.org/blog/2017/05/a-privacy-threat-model-for-the-people-of-seattle/

“[Introducing Threat Modeling For Seattlites](https://seattleprivacy.org/introducing-threat-modeling-for-seattlites/),” Jan Bultmann, July 9, 2017 https://seattleprivacy.org/introducing-threat-modeling-for-seattlites/

1. We’ll use “product” as shorthand for product, service, or combination thereof. [↑](#footnote-ref-1)
2. The word “bugs” here is a term of art, meaning anything a developer needs to work on. In this sense, bugs are often tracked for “we need a feature that does X” or “we need to write a test to see if this regresses.” [↑](#footnote-ref-2)
3. A dozen years ago, Window Synder wisely told me that this would be a problem. [↑](#footnote-ref-3)
4. We use XLSX rather than CSV because we use Excel’s “sheets” feature. [↑](#footnote-ref-4)