DESIGN AND ANALYSIS OF PRIVACY MECHANISMS FOR THE INTERNET OF THINGS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Kolbeinn Karlsson

August 2018

DESIGN AND ANALYSIS OF PRIVACY MECHANISMS FOR THE INTERNET OF THINGS

Kolbeinn Karlsson, Ph.D.

Cornell University 2018

The trend toward embedding internet-connected computers in objects and buildings is often referred to as the Internet of Things (IoT). The IoT has great potential to improve the efficiency of our homes and businesses. It also has the potential to severly undermine individual privacy. IoT devices generally report their states and sensor readings to their manufacturer at frequent intervals. This enables an unprecedented automated collection of information on ordinary citizens on a scale never-before possible. This information may seem trivial at first glance, but can in fact reveal a great deal of sensitive information.

Privacy mechanisms offer a potential solution to this problem. A privacy mechanism is an algorithm that enables one to enjoy the benefits of a service without revealing too much personal information. The IoT raises new privacy concerns that existing privacy mechanisms have not been designed to handle. These new problems require new methods to design privacy mechanisms. Engineering design is nothing without engineering analysis, however. We will also need new tools to quantify the privacy afforded by a mechanism and its effect on the utility of the service in question.

This dissertation presents three main contributions to the design and analysis of privacy mechanisms for the IoT. It begins with a technical, ethical, and legal examination of online advertisement and ad blocking. Online ads are simultaneously an economic driving force behind the increasing collection of

personal information as well as a mechanism enabling said collection. As such, they are integral to any examination of a privacy in the IoT.

The next contribution is in the field of location privacy. GPS-enabled IoT devices that users carry on their bodies, such as smartphones, enable broad collection of location information on individuals. Location information is some of the most sensitive information collected by the IoT and is therefore one of the most significant research problems in IoT privacy. The dissertation lays out the analysis and design of location privacy mechanisms that focuses on hiding the semantic information in location data rather than the geographical coordinates themselves.

The third contribution shows how the IoT can be used to enhance privacy instead of undermine it. It presents the design of a tamperproof log that can operate in a low-power, low-connectivity IoT network. The tamperproof log enables enforcement of an accountability policy that can be used to protect the privacy of medical records delivered to emergency physicians during disaster response.

BIOGRAPHICAL SKETCH

Kolbeinn Karlsson started his academic career studying Electrical and Computer Engineering at the University of Iceland in Reykjavik, Iceland in 2009. In 2011, he transferred to Cornell University in Ithaca, New York, where he graduated with a Bachelor of Science degree in Electrical and Computer Engineering in 2013. After working for a year as a computer vision researcher at Futurewei Technologies, he returned to Cornell University in 2014 to pursue a PhD, also in Electrical and Computer Engineering. This dissertation is the result of his work there from 2014 to 2018.

This dissertation is dedicated to my wife Katla, whose unwavering love and support got me through graduate school.

ACKNOWLEDGEMENTS

Many people contributed to my success in graduate school. I am particularly grateful to my advisor, Stephen Wicker. He gave me enough latitude to find my footing when I was starting out and gave me the encouragement I needed to follow my own research ideas as I progressed through the PhD program. His encouragement fostered independence and confidence in myself as a researcher, traits that will benefit me well beyond graduate school. I could not wish for a better mentor. My committee members, Aaron Wagner and Christoph Studer, have been extremely helpful too. Their feedback over the years helped shape my research towards greater clarity and rigour.

I would also like to thank Professors Hakim Weatherspoon and Robbert van Renesse, with whom I collaborated closely during my last year at Cornell. I cannot think them enough for the mentorship and guidance over the past year. During my collaborations with them, I also had many fruitful conversations about research with Doctors Weijia Song and Zhiming Shen and Professor Christina Delmitrou. A special thanks goes out to Hakim's student Danny Adams, who was instrumental in making Vegvisir a success.

Thanks to Andreas Unterweger and Professor Dominik Engel from the Salzburg University of Applied Sciences for their feedback on my location privacy research throughout our many conversations. I would also like to thank former M.Eng. students James Cassell, Edwin Ma, and Weitao Jiang for their contributions to the Android prototype of Vegvisir, and undergraduate researchers Ning Ning Sun and Sherbin Abraham for their feasibility study of ad hoc network protocols for Vegvisir.

I could not have gotten my work done without the extremely capable and helpful ECE administrative staff. A special thanks to Scott Coldren, Daniel

Richter, Cindy Vanostrand, Patricia Gonyea, and Mike Haldeman. I am equally grateful to Irwin and Joan Jacobs, the National Science Foundation, and the National Institue of Standards and Technology for their funding support. And I am eternally grateful to the Cornell Laboratory of Plasma Studies. They not only let me use their break room and drink their coffee for the past four years, but were also kind enough to include me in their lab lunches and recreational activies.

I might never have been admitted to the program if not for the recommendation letters written by Dr. Heather Yu and Professors David Delchamps and Tsuhan Chen. Each of them have also been valued mentors. I cut my teeth as an undergraduate researcher working with Tsuhan and later Heather in the field of computer vision. And David Delchamps, my undergraduate advisor and teacher, showed me the value of applying the rigour of mathematical theorems and proofs to engineering problems.

My friends in the ECE department made graduate school so much more enjoyable. My weekly Wednesday lunches with Nirmal Shende, Kia Khezeli, and Sam Lin were a highlight of my week. I am also so thankful to Sophia Rocco, Alex Ruyack, Mark Campell, Melissa White, Polina Alexeenko, Kia Khezeli, Nirmal Shende, and Levon Atoyan. They are not only good friends but also helped me rebuild the ECE graduate organization ECEGO and create a greater sense of community in the department. I also want to thank Alex Ivanov, Raphael Louca, and Jeffrey Mulligan for many good times over the past four years.

Last, but not least, I want to thank my family. The comfort and peace of mind from having a family that loves you unconditionally cannot easily be stated in words. Their unfailing belief in me spurred me on when times were tough and the frequent Skype calls with my brother, sister, parents, and in-laws made

me feel connected to Iceland and to my family while living thousands of miles away.

It is safe to say I would not have had the courage to even begin this journey if not for the love and support of my wife, Katla. She belived in me from the very start, and more importantly, also when my own belief in myself faltered. She is the smartest, kindest person I know and every day with her inspires me to become a better person.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

At its best, the IoT has the potential to create an integrated ecosystem that can respond to a spectrum of needs, increasing efficiency and opportunity, and empowering people through technology, and technology through intelligence. At its worst, the IoT can open a Pandora's Box of inappropriate and unsafe behavior, unintended consequences, and intrusiveness.

Vinton G. Cerf [13]

1.1 Introduction

The Internet of Things (IoT) is the name given to the trend toward integrating internet-connected computers into objects not traditionally connected to the Internet. The ubiquity of WiFi, advances in cellular data networks, and ever-cheaper and more powerful microprocessors have made it easier and more affordable to connect objects around us to the internet in novel ways. Modern cars have WiFi and a cellular data uplink to their manufacturer to report telemtry and download firmware upgrades. Modern "smart" electrical appliances, such as refrigerators, coffee makers, and vacuum cleaners, enable users to monitor and control them over the internet. Smart homes have door locks, windows, lights, environmental controls, and alarm systems all connected to the internet to enable increased energy-efficieny and remote monitoring. Smart speakers, such as Amazon Echo or Google Home, provide an alternative voice-activated interface to smart homes and to virtual assistants such as Apple's Siri and Amazon's Alexa. Smartphones, while on the boundary between IoT devices and

general-purpose computers, form an important class of IoT devices. All of these devices and more form a network of billions of ”things” that has drastically changed the structure of the internet.

The IoT has been the cause of a lot of excitement in industry, academia, and government. By most estimates, the number of IoT devices and revenue in the IoT market is projected to grow exponentially in the next decade. Both technology and mainstream press is flooded with articles on the wonderful potentials of the IoT. Dozens of industry consortiums have been founded around the IoT, numerous academic journals and conferences have been dedicated to it, and governments are investing heavily in IoT initiatives [88]. This excitement is not fully unwarranted. The IoT does indeed hold great potential. Internet-connected cars can download the latest map, traffic, and weather information to give drivers up-to-date navigational directions. They can send telemetry to the car manufacturer who in turn suggest preventative maintenance steps and automatically push software updates to the car computer. IoT-powered home automation not only allows users to conveniently control their appliances and environmental controls from their smartphones, but could potentially dynamically adjust heating, lights, and air conditioning to reduce power consumption and carbon footprints. And as new IoT devices enter the market, new possibilities for greater convenience and efficiency emerge.

As a side-effect of all these new services and capabilities, the IoT generates an unprecedented amount of data on individual users. Some of this information may seem trivial at first glance, but upon closer inspection it can reveal a great deal about the user. For example, if a company knows when and how you use your coffee maker, they can get a tiny picture of your daily life. They

might be able to infer when you wake up from the time you turn on your coffee maker in the morning. If you brew a large pot in the middle of the afternoon on a weekend, that could indicate you have company. The size of the daily pot brewed could indicate the number of coffee drinkers in the household and/or how much coffee you drink. This information can in turn be correlated to your age, gender, ethnicity, and socioeconomic background.

But this is just information from a single coffee maker. The information revealed is compounded by the number of IoT devices in a household. If the coffee maker data is combined with data from your smart electricity meter, your internet-connected thermostat, a smart home computer controlling doors and windows, air conditioning, faucets, and lights, soon enough it becomes feasible to infer just about everything that goes on in the home. As new IoT devices are added to the home, their data forms an increasingly complete picture of everything that transpires within a household. Outside of the house, wearables and smartphones can track user movements and activities. The location data alone can be incredibly revealing. It can reveal that you participated in a political demonstration, or a labor union meeting, or visited an HIV clinic. As we shall discuss further in chapter 3, given a detailed enough location trace on an individual, it becomes possible to reconstruct their life.

With all this data collection, the IoT opens up an entirely new realm of privacy problems. Never before has it been possible to collect, store, and analyze so much information about people. As the IoT grows to monitor and control more and more of our lives, it is able to create an ever-more detailed simulacrum of our private lives. The problem is compounded by the fact that this data is not necessarily siloed between different IoT device manufacturers. In fact, as we

shall see in chapter 2, there is a lively economy built around buying and selling personal information.

Digitized personal information is being used increasingly to make important decisions about our lives. Such data now affects whether people get into college, are offered a job, have their loan applications approved, or their insurance premiums increased [99]. Traditionally, the law has placed limits on what kind of information can be factored into such decisions, but law and ethics tend to be one step behind technological advances [127] and is not yet properly equipped to deal with the privacy problems of the IoT. Furthermore, the algorithms used to make these decisions carry with them the latent biases and prejudices of their designers and training data [89], often amplifying existing inequalities in our society [99].

Given the promise the IoT has for making our lives easier and more efficient, we would be remiss in simply doing away with the IoT altogether. Instead, we can use careful engineering to have our cake and eat it too. By transforming or censoring information prior to sending it to an IoT service provider, we could potentially still derive the aforementioned benefits from the IoT without revealing sensitive personal information. A technology that alters or censors personal information prior to sending it to IoT service providers will in this dissertation be referred to as a *privacy mechanism*. A privacy mechanism could add noise, reduce precision, add extraneous information, or even fully delete information prior to sending it over the network. Designing such mechanisms forms an important part of this dissertation.

In order to design good privacy mechanisms, we need to be able to tell good mechanisms from bad ones. In other words, we need tools to carry out the engi-

neering analysis of privacy mechanisms. Most importantly, we need quantitative measures the impact of privacy mechanism design decisions on the amount of information revealed as well as on the quality of service. This is the so-called privacy-utility trade-off and is challenging because both ”privacy” and ”quality of service” are inherently ambiguous concepts. We need to determine which approaches are appropriate for each given context. A significant part of this dissertation is devoted to the development of such privacy and quality-of-service metrics.

1.2 Dissertation Overview

This dissertation focuses on three main areas within IoT privacy: Mobile ad blocking, location privacy, and sensitive records management. Chapter 2 explores the topic mobile advertising and ad blocking, a topic of pivotal importance to IoT privacy. The online advertising economy drives the demand for consumer personal data, which in turn is the impetus for the intense commercial interest in connecting everything to the internet. In fact, sometimes companies don't even know what additional value an embedded internet-connected computer brings to the customer [56].

IoT service providers undoubdetly know what value IoT devices bring to them. Personal data is used to create a detailed marketing profile on individuals. Once these profiles are created, more data is then needed to track the exposure and response of those individuals to specific advertising campaigns. Mobile devices such as smartphones have become the most important arena where this information collections plays out. Facebook, Inc. is one of the largest

online advertising companies in the world and reportedly makes 91% of its revenue from mobile advertising [122]. As such, online ads have ramifications beyond the IoT, but the IoT remains a central battlefield in the arms race between advertisers and adblockers.

Mobile advertising is interesting because it is not only a cause of personal data collection, it is also one of the primary mechanisms by which it is done. Online ads are the primary tool by which users' browsing habits are tracked across the web. We will explore how ads are used to track user behavior as well as how ad blockers, an important type of privacy mechanism, work. As ad blocking is a controversial technique, we will also provide a brief legal and ethical analysis of the use of ad blockers.

In chapter 3, we move on to one of the most intrusive forms of personal information collection in the IoT: location information. I will review existing mechanisms and approaches to analysis in the field of location privacy and go on to propose a different approach that more closely reflects the privacy needs of a typical user. The central idea is to protect the semantics associated with location information rather than the geographical coordinates themselves. Protecting semantics is more complicated than simple coordinates, but I will show how some of the tools developed to measure physical location privacy can be extended to semantic location privacy and show how to use them to evaluate the privacy-utility tradeoff of mechanisms. The chapter concludes with a presentation of a privacy mechanism of my own design and its analysis using the tools developed in the chapter.

In chapter 4, we will see that while the IoT creates a host of privacy problems, it also has the potential to extenuate them. In this chapter, we look at the prob-

lem of providing access to sensitive personal records such as health records to emergency first responders during disaster response. During disaster response, existing IT and communications infrastructure may be rendered unavailable, but do to the prevalence of smartphones, emergency first responders bring with them a considerable amount of computing power and communication capabilities just by the virtue of having their phones on them. We will present Vegvisir, a blockchain-based system operating across low-power IoT devices communicating in an opportunistic fashion. Vegvisir can enable emergency medical responders operating outside of traditional communication infrastructure to obtain access to the medical records they need without sacrificing the privacy of said records.

Finally, chapter 5 summarizes the work and conclusions presented in the dissertation. We will also point to future research directions within the areas covered in this dissertation as well as within privacy problems in the IoT as a whole.

CHAPTER 2

INTERNET ADVERTISING, TRACKING, AND AD-BLOCKING

Every time you block an ad, what youre really blocking is food from entering a childs mouth. [105]

In reality, ad blockers are one of the few tools that we as users have if we want to push back against the perverse design logic that has cannibalized the soul of the Web. [133]

2.1 Introduction

In fall 2015 Apple introduced a content blocking extension point into its Safari mobile browser, providing a hook for software that prevents advertisements from being loaded when web pages are rendered [63]. As it turns out, large numbers of people wanted to do just that [132]. Ad blockers had been available for some time, but their potential use in the worlds most popular mobile browser heightened their saliency and brought the debate over their use, a debate sometimes serious and nuanced, but often frivolous, into the mainstream media [33].To put the issue into perspective, consider the following provided by PageFair, a leading provider of counter ad block solutions to web publishers, in its 2015 [102] and 2016 [101] reports on ad blocking:

- Ad blocking was estimated to have cost publishers nearly $22 billion during 2015.

- As of November 2016, at least 309 million people are blocking advertising on their smartphones.
- 298 million of these people use an ad blocking browser, more than twice the number using blocking browsers in 2015.
- Ad blocking is particularly popular in emerging markets, with the largest number of active monthly users in China, India, and Indonesia. The United States is in ninth place.

In its 2016 report PageFair made the following prediction:

> *Mobile ad blocking is a serious threat to the future of media and journalism in emerging markets, where people are coming online for the first time via relatively expensive or slow mobile connections. Usage in western economies is likely to grow as more manufacturers and browsers start to include ad blocking as a feature.* [101]

Given the amount of money involved in advertising, one might expect a certain amount of invective on the subject of ad blocking. One would be correct. Ad blocking has been referred to as evil and as a form of theft [107]. Ad Age, an advertising industry trade magazine, accused ad blockers of being exploitative, extortionate, and anti-democratic, all within the space of a single sentence: As abetted by for-profit technology companies, ad blocking is robbery, plain and simple an extortionist scheme that exploits consumer disaffection and risks distorting the economics of democratic capitalism. [109]

Randall Rothenberg, president and chief executive officer for the Interactive Advertising Bureau accuses ad blocking profiteers of stealing from publishers,

subverting freedom of the press, operating a business model predicated on censorship of content, and ultimately forcing consumers to pay more money for less and less diverse information [62].

On the other side of the debate, many have pointed to the ads themselves as fostering needless consumption while being tasteless, intrusive, and evil (this word occurs a lot in these discussions), while suggesting that the advertising industry brought ad blocking upon itself [3].

There are purely technical issues as well. The technology that allows Internet advertisers to better target potential consumers slows the loading of web pages and places a significant burden on wireless cellular links, a burden that is usually funded by unwilling users. The ad-blocking software provider Shine, an Israeli startup that began life in 2011 as an anti-virus software developer, estimates that advertising consumes between 10% and50% of user data plans, depending on user location. A typical mobile gaming app with advertising was found to consume 5Mbytes over a five-minute session, but only 50 Kbytes with ad blocking in place [100].

Shine produces ad-blocking software that can be incorporated into cellular data centers. In June 2016 the UK cellular service provider Three became the first service provider to conduct trials using this software to block ads on cellular data connections [65]. Given that marketers are expected to have spent over $100 billion on mobile ads in 2016 [57], the response is expected to be extreme.

In this chapter we explore how advertising networks and ad blockers work. We further consider how ad blockers are subverted, and whether they are ethical. The ethical analysis yields mixed results, but it does, however, suggest a

solution that empowers users, allowing them to select the types of advertising they see and how often they see them.

The work in this chapter has previously been published in Communications of the ACM [131].

2.2 The Technology of Ad Networks and Ad Blockers

Web browsers request a web page from a server by sending an HTTP GET command to the appropriate Internet host. The host responds with HTML code that the web browser uses to render the desired page and present it to the user. This much is both simple and ubiquitous, but the details, particularly when advertising is involved, are much more complicated.

Suppose that the web browser requests a page from a content publisher that supports his or her work through advertising (this is represented in the figure below by link 1). Most publishers do not generate their own advertising content, so they will embed requests for advertising into the HTML files that they send to requesting users (link 2). When the requesting host attempts to render the HTML file, it will generate requests for advertisements from an ad exchange. The ad exchange, as shown below, sits at the center of a network consisting of supply side and demand side entities. The supply side entities provide information about the user, while the demand side entities provide advertising in response to requests from publishers.

The HTML code provided by the publisher directs the host to a supply side platform (SSP link 3 above). The request sent to the SSP includes a cookie - a

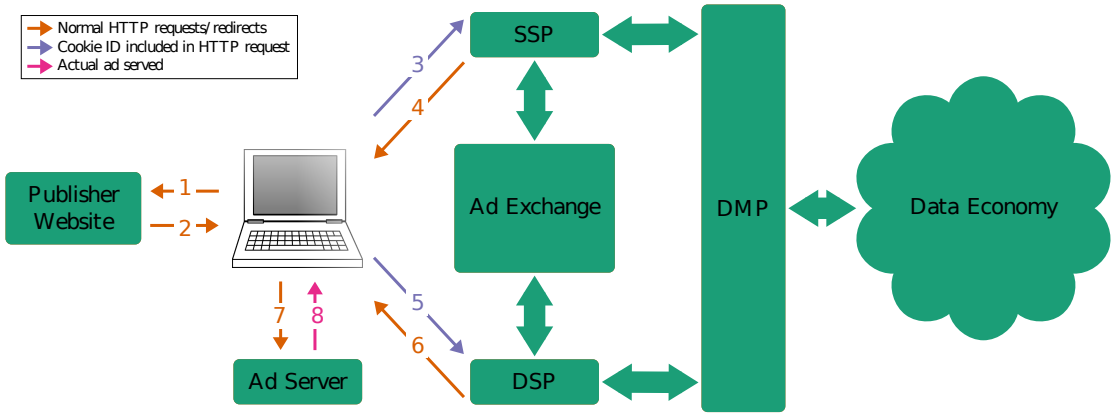


Figure 2.1: Entities involved in delivering a web page ad.

small string of information that was previously stored by the SSP on the users computer. The cookie enables the SSP to craft a response that is specifically tailored to the requesting user. In this case, the cookie will include a user ID that the ad exchange can use to coordinate bidding for an advertisement. The ad exchange forwards the user ID and any other information that it may have about the requesting user to one or more demand side platforms (DSPs) that place bids on behalf of advertisers for the opportunity to display their ads. Through a process known as cookie syncing, the DSPs are able to match the SSP cookie ID to a user profile, which is often stored and managed by a separate entity called a data management platform (DMP).

As multiple SSPs and DSPs can use the same DMP, the DMP may link a wide range of user IDs to the same person. This enables all interested parties (other than the user) to exchange information on the user and form a more complete picture of that users browsing history. First party websites may also participate in the process, providing yet more user information. For example, if a user supplies an email address to a website to sign up for their newsletter, the email address can be linked at the DMP to the cookie IDs associated

with that user. If the user provides a name and address to a website, that information may also be linked to the cookie IDs. The DMP may take this a step further by including information inferred from the user's social media activity, purchase history on various sites, search history, and emails. Finally, the DMP may have access to data gathered offline. Data aggregators are known to collect data from publicly available records, including licensing records (e.g. licenses for doctors, lawyers, pilots, or hunting and fishing licenses), voter registration databases, court records, and DMV records, as well as buying data from commercial sources including brick-and-mortar store purchase histories and transaction information from financial services companies [29]. Data aggregators also buy and sell information from each other. This whole system of transactions is often referred to as the data economy. Through this data economy, the DMP is able to build a strikingly detailed simulacrum of an individual consumer, a simulacrum whose accuracy drives the advertisers return on investment, and whose inaccuracy may drive the consumer to distraction.

If a DSP determines that the user profile fits its target audience, it places a bid for advertising space on the web page being rendered by the host computer. The ad exchange selects an ad from amongst the bidding DSPs; a Vickrey auction is generally used, where the highest offer is selected and the amount paid is that offered by the second highest bidder. The winning DSP provides a URL for retrieving the ad. In an impression based system, an agency ad server determines whether the ad is actually downloaded, and pays the publisher accordingly, with the cost per thousand impressions (CPM) being the most widely used statistic in Internet advertising [41]. All of this happens within tens of milliseconds, though the actual loading of the winning ad into the users browser may take far longer depending on the bandwidth of the user connection and the

size of the ad.

As a result of this process, the publisher of the content often has limited control over the safety, quality, or tastefulness of the ad seen by the content consumer. A publisher may, for example, be able to prevent advertisements from a particular advertiser or class of advertisers, but she may not be able to exercise finer control.

Ad blockers can use several methods to disrupt the above-described process and thus prevent ads from being displayed. Many prominent ad blockers, such as Adblock Plus and its variants, block ads by preventing the browser from sending HTTP requests to certain URLs. The URL blacklist for a given blocker is often a crowd sourced effort, such as EasyList, the default blacklist used in Adblock Plus. EasyList is probably the most widely used blacklist; the number of EasyList downloads was used by PageFair and Adobe to estimate the prevalence of ad blocking in their 2015 joint report [102].

While URL blacklisting appears to be the most common method of ad blocking, the Electronic Frontier Foundation's Privacy Badger takes a different approach [45], attempting to learn which domains and sites are tracking a user and blocking the ones that do. It detects behavior such as the use of uniquely identifying cookies, canvas fingerprinting, and the appearance of the same third party site at multiple domains. As such, it blocks very few domains at first, but the more it is used, the more it learns to block. It should be noted that Privacy Badger aims to prevent tracking, not ads; but since the two are intimately connected, it often serves both purposes.

A third ad blocking method blocks website elements fitting certain patterns;

for example, it could look for the ”iframe” HTML tag and check to see if it contains text strings like ”Sponsored” or links to a URL with the word ”ad” in it. This method can block advertising served by the web site itself as opposed to just third parties, advertising that includes ads embedded in search results and social media feeds. This content filtering can happen at the client or at an intermediate proxy. Some ad blockers use a root certificate to redirect browser requests to a VPN or proxy that removes ad content using the above-mentioned methods before forwarding the HTML code to the browser. This approach can block ads for mobile apps as well as browsers, but it comes with the risks associated with having third parties interfere with browser traffic, risks that include the classic man-in-the-middle hacking attack. Apple recently removed several ad-blocking apps from its app store on this basis [52].

Publishers will sometimes try to circumvent attempts at ad blocking. Anti-ad blocking usually works by serving a fake ad in some way and verifying that it has been loaded or displayed. If it fails to load, the site stops displaying the primary content or refuses to load it in the first place. For example, a site can contain an iframe ostentatiously marked as ”Advertising” and then use JavaScript to see if it was displayed a . If an ad-blocker is detected, the site doesn't provide the primary content. Similarly, the browser can be directed to load a JavaScript with a name, such as ”ads.js,” that can be found in common ad blocker filter rules and check to see if it is run. Aside from trying to explicitly detect ad blockers, ad networks can obfuscate the URLs of their ads, such as by using IP addresses instead of domain names.

Ad blockers can often adapt, circumventing new anti-ad blocking mechanisms. Facebook recently announced that it would prevent ad blocking [85],

only to have Adblock Plus announce it had found a way to defeat Facebook's prevention technique a few days later [58]. This is but one example of the evolving arms race between publishers and ad blockers [95].

Though the initial motivation for ad blocking may be annoying ads or tracking, increased computer security is a major side benefit. Online ads are usually pieces of code as opposed to static images or text. The end result of the ad auction process described above is that the user's browser is redirected to a URL of the advertisers choosing. The retrieved object may take the form of JavaScript, Flash, or even Java code. Vulnerabilities in these frameworks can be used to execute malicious code on the client machine without the user noticing anything out of the ordinary. Even though browser support for Java and Adobe Flash is being phased out [60], vulnerabilities in these frameworks are still being exploited. Java exploits are on the decline, but Flash vulnerabilities are still some of the most common vehicles for malvertising [124]. While ad networks have measures in place to detect malvertising, there are ways to circumvent and avoid detection, at least temporarily, such as serving a legitimate ad until the ad network has approved the ad, only serving malvertising every 10th or 20th time, and not serving malvertising to certain IP addresses [74]. Even large and reputable websites have been known to accidentally serve malvertising, making malvertising a potential problem for every Internet user.

2.3 Is ad blocking a breach of contract?

It would take a long law review article, and one written by another set of authors, to properly address the legality of ad blocking. We do, however, wish to

address the oft-cited argument that the provision of free content that contains ads is done under an implicit contract [91]. Under this contract, the consumer is provided with free content in return for the users agreement to view advertisements. This is not a new argument, as it has been applied by network executives to broadcast television for many years, sometimes in a very extreme form. In 2002 Jamie Kellner, then CEO of Turner Broadcasting, suggested that any systematic practice of using the bathroom during commercials was stealing [44].

In the United States, the legal concept underlying this argument is the implied-in-fact contract. The law is summarized as follows:

> *To establish the existence of an implied in fact contract, it is necessary to show: an unambiguous offer, unambiguous acceptance, mutual intent to be bound, and consideration. However, these elements may be established by the conduct of the parties rather than through express written or oral agreements. [64]*

As an example, suppose that you agree to wash your neighbors car once a week. You receive payment for each of the first six weeks, but upon washing his car the seventh time, your neighbor refuses to pay because there was no written agreement. Most courts would agree that there was an implied-in-fact contract as evidenced by the conduct of the parties for the first six weeks. Your neighbor has to pay. Now consider a real example in which it was found that there was no implied-in-fact contract. In 1917 the United States leased a pier from the Baltimore and Ohio railroad for the purpose of handling supplies destined for the war in Europe. An earlier fire was believed to have been an act of sabotage,

so soldiers were deployed to guard the pier and surrounding equipment. The weather was cold, and the troop commander often complained about the tents in which his men were forced to live. A railroad official offered to build temporary barracks. Though there was never any discussion of compensation, the barracks were built. The railroad later sued to recover the cost of the construction, arguing that there had been an implied-in-fact contract. In what became the 1923 case of Baltimore & Ohio R. Co. v. United States [1], the Supreme Court disagreed. The Court stated that an implied agreement required a meeting of minds inferred, as a fact, from conduct of the parties in the light of surrounding circumstances. The Court found that there had been no such meeting of minds, as the railroad company never intimated that it would expect payment from the government.

It follows that there are several reasons that the alleged quid pro quo of viewing ads in return for free Internet content fails to rise to the level of an implied-in-fact contract. First, as with the Baltimore & Ohio case, there was no unambiguous offer. The Internet content consumer is rarely told precisely what is going to be loaded into his or her web browser, and what is expected in return. Content consumers suffer the embedding of ads and, on occasion, trackers and other forms of spyware into their web browsers without receiving any notice from the content provider whatsoever. In fact, as we have seen, the content provider may not know what is being injected into the consumers browser.

Second, the alleged agreement fails to satisfy the unambiguous acceptance element. Unlike the lawn-mowing example, there is no prior conduct that indicates a general understanding that an agreement is in place. The popularity of ad blockers [102, 101, 94] indicates that most consumers do not want to see the

ads, and clearly have not agreed to do so. A Reuters survey provides further evidence, indicating that even those who do not employ ad blockers are ignoring or avoiding the ads:

> *More generally, a third or more (39% in the UK and 30% in the US) say they ignore ads. Around three in ten (31%/29%) say they actively avoid sites where ads interfere with the content. [94]*

2.4 Are Ad Blockers Unethical?

In After Virtue Alasdair MacIntyre describes the breakdown in ethical argument that occurs when the foundations for ethical systems are cut away, leaving proponents of differing perspectives to argue past each other without any basis for decisive engagement [84]. Ad blocking provides a canonical example, as we have one group arguing for individual rights (the right to receive payment for ones effort in providing content), while the other group argues for the general welfare (an Internet devoid of continual distraction caused by tasteless advertising). It is not clear how the two arguments can be reconciled, or how one can clearly overcome the other. We suggest that a solution lies in a technologically-mediated meeting of minds, but before we consider the solution, we offer a more detailed account of the ethical arguments.

The utilitarian approach, first propounded by Jeremy Bentham and John Stuart Mill in the late eighteenth and early nineteenth centuries, is based on the familiar precept that it is the greatest happiness of the greatest number that is the measure of right and wrong [10]. In what follows we will consider act utilitar-

ianism, which focuses on the consequences of individual actions. We will also substitute well-being for happiness to counter some of the more obvious criticisms of utilitarianism. Does the use of ad blockers create the greatest well-being for the greatest number? Those affected by the decision to block ads include the following:

- Ad Blocking Users
- Ad Viewing Users
- Content Generators
- Content Publishers
- Advertisers

Should users choose to employ ad blockers, the following will arguably result:

- The ad blocking users will see fewer advertisements.
- The content generators will receive less revenue per reading user.
- The content publishers will receive less revenue per reading user.
- Advertisers will seek other venues for their advertising dollars.
- Some content generators will stop generating content.
- Some content publishers will stop publishing content.
- Some content publishers will publish content of lower quality.
- **There will be less free content available to all users on the Internet, and the content that remains freely available will, in some cases, be of reduced quality.**

It is important to provide some context for the suggestion that the quality of online content will be diminished by a general acceptance of ad blocking. Newspaper journalism was in decline well before the advent of ad blocking, or even the advent of the Internet, primarily because of the failure of its core business model [86]. The business model was that of a quasi-monopoly: competition was limited, so that a local paper could charge higher prices for advertising, and then use the revenue to maintain reporters across the world. In essence, the local Wal-Mart paid for the Baghdad bureau through its advertising dollars. The limit on competition was due to a fact of technology: printing presses were very expensive to operate and maintain, so all but the largest municipalities could only sustain one or two (print) newspapers at any given point in time [114]. In a pre-Internet world, the papers acted as an intermediary between advertisers and consumers, charging both for the opportunity to communicate. In a multi-newspaper market, the equilibrium was often unstable; a notable scoop could send more advertising dollars to the scooping paper, allowing the scooper to grow (literally) fatter and more attractive to the buyers.

The unraveling of this relationship began with the television era and the movement of affluent readers from the inner city to the suburbs. National and retail advertisers moved their dollars to television, and newspapers came to depend more on classified ads [37]. With the advent of the Internet in general and Craigs List in particular (founded in 1995), classified advertising revenue also began to leave the newspapers balance sheets. By 2010 the newspaper industry was in deep decline, with many major players facing bankruptcy (e.g. the Tribune Company in 2008), and others left to cope with dramatically reduced staffs.

The consequences of a general use of ad blockers may thus be characterized as a further reduction in the quality of free online content through the departure of some Internet content generators and publishers to other ways of making a living. For large numbers of consumers these are apparently acceptable outcomes given what they avoid: the problems associated with spyware and the relentless distraction of advertising. There is also evidence that Internet readers do not greatly value what they are reading; given the choice between paying for the content and losing it, most prefer the latter. The aforementioned Reuters survey found that only 10% of online users appeared to be willing to pay for once-free news content:

> *After a sharp upturn in 201213 when a large number of paywalls were introduced our data show very little change in the absolute number of people paying for digital news over the past year. In most countries the number paying for any news is hovering around 10% of online users and in some cases less than that.* [94]

If Internet readers and users of ad blockers are rational actors who are making decisions based on their individual well-being, and as the readers outnumber the writers and advertisers, one may conclude that the use of ad blockers provides the greatest wellbeing for the greatest number. From a utilitarian perspective, ad blocking is ethical; the content providers should look for a better business model. The counter-argument is ready at hand: this analysis clearly does not take into account all stakeholders; the content generators and publishers, for example, would almost certainly not be pleased with the consequences of this utilitarian calculus. This is an example of a key criticism of utilitarianism; namely, that in emphasizing aggregate well being, some individuals may be left

in far worse condition than before.

A deontic analysis avoids this particular problem. As is well-known, Kant suggested in his Groundwork of the Metaphysic of Morals that there is a single primary moral obligation, which he referred to as the categorical imperative (CI). Kant offered several formulations of CI, including one that sounds very much like the golden rule: act so that you use humanity, as much in your own person as in the person of every other, always at the same time as end and never merely as means" [66]. Ad blocking readers arguably do not satisfy this formulation they treat the content generators as means rather than an end in themselves, taking their work product without respecting their efforts to make a living. It appears that Kant is on the side of the advertisers, while Bentham favors the general reader.

Contractualism, an ethical theory related to Kants deontological approach [103], more clearly takes into account all interested parties, while pointing to a potential solution. In What We Owe Each Other, T. M. Scanlon offers the following ethical rule for action [111] (emphasis added):

> *An act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behavior* **that no one could reasonably reject as a basis for informed, unforced, general agreement.**

In establishing rules for behavior, Scanlon suggests that we must consider the perspectives of all stakeholders, and define a basis for informed general agreement. This would require communication between all stakeholders, something that is sorely lacking in the context of online advertising. We will return to

this point when we consider possible solutions. The third and final approach to be considered shifts the balance of the argument in favor of the general reader, but on a far firmer basis than the arguments of Bentham et al. Aretaic, or virtue ethics, emphasizes virtues of mind and character [31, 33]. Virtue ethics originated with Aristotles Nicomachean Ethics and his notion that the ultimate aim (telos) of an individual is to live a virtuous life. A virtuous life is a life lived according to reason, where decisions are based on a set of values held dear by the individual. Virtue ethics thus involve the questions of what is desirable, good or morally worthwhile in life? What values should we pursue for ourselves and others? [49]

Virtue ethics has enjoyed a recent resurgence, both in philosophy departments and in schools of technology. With regard to the latter, value-based design practices have been developed based on various lists of fundamental human values. For example, in Ethical IT Innovation, Sarah Spiekermann points to both Aristotle and Maslow while concluding that technical design must be based on an understanding that knowledge, freedom, and autonomy are preconditions for human growth, self-esteem, friendship and self-actualization [114]. At best, the design of advertising technology shows little concern for knowledge, freedom, and autonomy of consumers. At worst, advertising technology actively works to subvert these values. This subversion can be seen through the lens of the attention economy, a term coined by Herbert Simon to capture the finite nature of the individuals attention in the face of a seemingly infinite amount of information [118]. The attention economy is reflected in advertisers insertion of themselves into virtually all personal interactions in everyday life, ranging from highway billboards to doctors offices to the bottoms of the trays at airport security. Writing for the Practical Ethics blog of Oxford University, James Williams

argues that the resulting distractions are more than an annoyance, they keep us from living the lives we want to live:

> *In the short term, distractions can keep us from doing the things we want to do. In the longer term, however, they can accumulate and keep us from living the lives we want to live, or, even worse, undermine our capacities for reflection and self-regulation, making it harder, in the words of Harry Frankfurt, to want what we want to want. Thus there are deep ethical implications lurking here for freedom, wellbeing, and even the integrity of the self.* [133]

From a virtue ethics standpoint, it follows that the design of Internet advertising technology is itself unethical in that it works against the human project of self-creation. Ad blockers are thus not only ethical, but are literally a matter of self-defense. Quoting the Practical Ethics blog once again:

> *In reality, ad blockers are one of the few tools that we as users have if we want to push back against the perverse design logic that has cannibalized the soul of the Web.* [133]

2.5 Solutions and Conclusions

The advertising delivery systems described in this chapter are the antithesis of value-based design. The values that Sarah Spiekermann and others point to as a foundation for virtue-based design knowledge, freedom, and autonomy

are precisely the values that online advertising systems most systematically undermine. Internet advertisers exchange information about users without their knowledge or control, using that information to manipulate users into behavior they might not otherwise have exhibited. This summary may seem harsh, and some may argue that advertisers would happily engage in more ethical behavior if better channels of communication were provided to interested consumers. A solution beneficial to all may lie in a virtue-based redesign. Such a redesign would embed T. M. Scanlons suggestion that there be an informed, unforced, general agreement among all parties. The agreement would be based on a system that provides revenue for content generators and connects advertisers to interested consumers while reducing the deleterious impact of the current system of advertising on the reading public. The key step lies in empowering the reading/consuming public letting them choose whether they will download ads, and if so, what type of ads. Should a reader choose not to download ads, he or she should be given the opportunity to pay for ad-free content. The supporting technology for such a system already exists in the current ad networks.

Recall that the current scheme of directed Internet advertising relies on the use of cookies stored on user machines. These cookies are sent to service side and demand side platforms to obtain directed advertising for insertion into content initially requested by a user. Suppose that the cookies are replaced by information explicitly provided by the user that indicates buying habits and interest in specific consumer goods. The data management platform (DMP) would request, coordinate and update this user-supplied information as necessary. Rather than inferring potential sales from browsing habits, advertising networks could make advertising bidding decisions based on the clearly expressed desires of potential consumers. Such a system would increase the agency of the

browsing user, while potentially increasing return on investment for advertisers.

Such a solution will require careful design and far more communication between stakeholders than currently takes place, but it offers the potential for clearly informing readers of their options, options upon which they can exercise rational choice in pursuit of their own individual goals. We hope that advertisers see this as an opportunity.

We have argued that ad blocking is not a violation of an existing contract (at least in U.S. law). This does not mean that ad blocking is beyond the reach of earnest lobbyists and subsequent legislation. One might expect, however, that such legislation would not be very popular with the general public. We hope that the above-suggested agreement takes form before the battle between advertisers and ad blockers escalates any further.

CHAPTER 3

SEMANTIC LOCATION PRIVACY

GPS information provides an intimate window into a person's life, revealing not only his particular movements, but through them his familial, political, professional, religious, and sexual associations. The location records hold for many Americans the privacies of life.

Chief Justice John Roberts [2]

3.1 Introduction

Location-based services (LBS) are some of the most popular innovations to come out of mobile computing and the Internet of Things. Turn-by-turn navigation apps like Apple Maps and Google Maps have rendered paper maps all but obsolete. Location-based ride-hailing services like Uber and Lyft have massively distrupted the age-old taxi industry. And location-based gaming took the world by storm when Pokemon Go was released in 2016 [134]. New location-based services are introduced all the time and we have undoubtedly not yet seen the full potential of location-based services.

Despite all their promise, location-based services also raise some of the most serious questions about privacy. Fine-grained location information can reveal an extraordinary amount of information about a person. To start with, it can very quickly reveal a person's home and work location. By combining the route taken to work with the speed traveled, it becomes possible to figure out the mode of transportation used. Visits to shops and clubs reveal buying habits and

hobbies, while visits to places of worship, medical clinics, and labor union halls can reveal sensitive personal attributes. If an individual's location trace is combined with that of others, it becomes possible to infer a social network between the people in the dataset. It could reveal one's friendships and acquaintances, and even romantic relationships.

The simplest types of location inferences are what I call *first-order* inferences. They are the result of *reverse geocoding*, the process of converting geographical coordinates to a place name or street address. Reverse geocoding a location trace produces a list of places a person has visited and potentially the duration and time of day of their visit. This list alone can provide quite the detailed picture of a person's life in and by itself. But first-order attributes are often correlated with ethnicity, income level, education, and other sensitive attributes, providing an even more intrusive picture of the individual. Combined, this information ends up revealing far more the user might have intended when they downloaded that location-based game or restaurant recommendation app.

This is where privacy mechanisms come in. It is possible to reduce the amount of information revealed in a variety of ways. We could add noise to the coordinates, or reduce their precision, or provide false coordinates along with the real ones. There is in fact, no shortage of location privacy mechanisms, as we shall see in the coming sections. The real question is how to evaluate and compare them. We need to quantitatively measure the amount of personal information contained in location information and to measure the effect of privacy mechanisms on the quality of service. Crucially, we need to understand how to strike a good trade-off between privacy and quality of service.

How one quantifies location privacy is dependent on what information is

considered sensitive. Early schemes such as spatial cloaking [55] and mix zones [12] assume that the sensitive information is a person's identity. While there certainly are cases where a person wants to hide their true identity from an LBS provider, it is often necessary or desirable to provide that information freely. Location-based services often have a social networking component to them, making it desirable for people associate their presence on the network to their real-world identity. Other services require personally identifying information for payment purposes. In such cases where the user has already provided their identity to the LBS, it becomes more important to limit the information revealed aside from identity.

In later work, the dominant approach became to protect a user's physical location coordinates. The two main approaches, based on Bayesian decision theory and differential privacy [38], respectively, differ mostly in their assumptions about the prior knowledge of the adversary. Protecting physical location can be important in some cases, such as military applications, but I claim that for your average smartphone user it is an indirect measure of privacy. The real intrusion to privacy happens after physical coordinates have been reverse geocoded to place names. Unlike geographical coordinates, a place names carries *semantics*. The semantics of the place is what reveals information about the user. I have therefore elected to take a different approach to location privacy, one where the focus is on protecting the semantics rather than the coordinates.

A place name, sometimes also referred to as a point of interest (POI), carries several layers of semantics. If I reveal that I am at "Starbucks on the Commons in Ithaca, NY", then I have revealed the fact that I am at a specific Starbucks coffee house. But it would also be true to say that I am at a "chain coffee house",

or even just "coffee house". While I acknowledge that there are complexities inherent in talking about the semantics associated with a place, I will in this thesis make the simplifying assumption that the place name is the only semantics associated with the place. The goal then becomes to prevent or minimize the ability of the adversary to infer the semantic location (i.e. the place name) of the user.

This chapter is dedicated to developing tools to analyze the location privacy and quality of service tradeoffs of privacy mechanisms under the semantic location privacy framework. We will show how to adapt both the Bayesian and Differential Privacy approaches from physical location privacy to semantic location privacy and we will develop both empirical and theoretical tools to evaluate the privacy-utility tradeoff. We will also present a location privacy mechanism of our own design, specifically created to protect semantic location information, and analyze it using the tools developed in the chapter.

3.1.1 Chapter overview

We begin the chapter by laying down notation and terminology and introducing the main families of location privacy mechanisms in section 3.2. Then we move on to a review of approaches to analyze physical location privacy in section 3.3. In section 3.4, we introduce the approach that I focus on, semantic location privacy. Section 3.5 provides an empirical Bayesian analysis of semantic location privacy mechanisms, while we derive theoretical results grounded in stochastic geometry in section 3.6. In section 3.7, we show how to adapt differential privacy to semantic location privacy, and in section 3.8 we present a

privacy mechanism designed to do well under both a Bayesian framework and a differential privacy framework.

3.2 Location Privacy Mechanisms

We begin by laying down the basic terminology and notation to be used in this chapter as well as to explore the basic families of location privacy mechanisms: adding noise, quantization, and adding dummy locations.

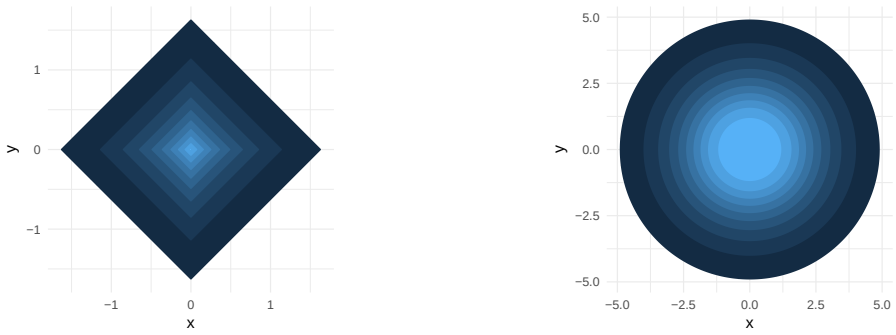
3.2.1 Terminology and Notation

Let us clarify a number of terms that we will use throughout this chapter. A *location report* is a tuple (s, t), where $s \in \mathcal{S}$ is a location, and $t \in \mathcal{T}$ is a timestamp. In most cases, $\mathcal{S}$ is a subset of $\mathbb{R}^2$, if location is assumed continuous, or $\mathbb{Z}^2$, if location is taken to be discrete. A precise definition of $\mathcal{T}$ is not important for our purposes as long as it is a totally ordered set. A *location trace* is a finite sequence of location reports associated with a single user, $r_1, ..., r_n$, usually in increasing order by timestamp.

Location privacy mechanisms are usually denoted as K and map, perhaps probabilistically, from the set of real locations $\mathcal{X}$ to the set of possible reported locations $\mathcal{Y}$. In other words, $K : \mathcal{X} \times \Omega \to \mathcal{Y}$, where Ω is the sample space. In many cases we will have $\mathcal{X} = \mathcal{Y}$.

3.2.2 Adding noise

Adding noise is a simple and well-understood technique for increasing the uncertainty around a value. In its simplest form, it can consist of adding a zero-mean, Gaussian random value to the coordinates before reporting them to the



(a) (b)

Figure 3.1: Contour plots of the naive vs. the planar construction of a 2D Laplace distribution. The planar Laplace distribution has circular symmetry, a desirable property that makes it easier to work with.

LBS provider. If $X \in \mathcal{X} \subseteq \mathbb{R}^2$, then we add $W \sim \mathcal{N}(0, \sigma^2 I)$ to it before reporting it:

$$Y = X + W$$

Laplacian noise is also a popular choice. However, the Laplacian distribution doesn't have as natural a generalization to multiple dimensions as the normal distribution does. You could construct a 2D Laplacian distribution from two independent 1D Laplacian random variables. Instead, it is common to use the planar Laplace distribution (see eqn. 3.1), where the angle is uniformly distributed over $[0, 2\pi)$ and the distance from the origin is given by the Laplace distribution (see figure 3.1).

3.2.3 Quantization

One of the problems with adding noise to location coordinates is that noise can be filtered out given enough location reports. For example, if the adversary controls how often the user submits a location report, they can reduce the effective noise arbitrarily by increasing the report frequency. A way to get around that is to reduce precision, or quantize coordinates, instead of adding noise to them. In the simplest case, coordinates are quantized with a quantization parameter Δ. If the real locations $\mathcal{X}$ are a subset of $\mathbb{R}^2$, then we apply a quantization function Q to each coordinate:

$$Q_\Delta(x) = \Delta \left\lfloor \frac{x}{\Delta} + \frac{1}{2} \right\rfloor$$

3.2.4 Dummy locations

The third major family of location privacy mechanism adds extraneous information to the location reports. One way that could happen is by reporting multiple locations for each location report. In other words, a dummy location privacy mechanism maps a real location $x \in \mathcal{X}$ to a set of real locations, so the range $\mathcal{Y}$ is a subset of the powerset of $\mathcal{X}$, denoted $2^{\mathcal{X}}$: $K : \mathcal{X} \times \Omega \to 2^{\mathcal{X}}$.

3.3 Analysis of physical location privacy

In this section, we will review the tools and techniques developed to analyze physical location privacy. The most important part of the engineering analysis of location privacy mechanisms is to quantify its effect on location privacy. We will present a number of metrics and evaluate their strengths and shortcomings in the coming sections. Measuring privacy is only one side of the coin, however. The final section of this chapter will be dedicated to reviewing the measures of quality of service or utility that have been proposed in the context of physical location privacy. For the sake of completeness, we will also review k-anonymity, an approach to privacy that was prevalent in the field of location privacy in its inception.

3.3.1 k-anonymity

The concept of k-anonymity originated as a technique to anonymize sensitive data such as medical or financial information, prior to publishing it [123]. The naive approach to anonymizing sensitive data is to simply remove all explicit identifiers such as names or social security numbers. The problem is that datasets often contain a number of *quasi-identifiers*. Quasi-identifiers are attributes that by themselves do not uniquely identify a person, but can be combined with other quasi-identifiers to uniquely identify an individual. The author of [123] found that 87% of the United States population could be uniquely identified by a combination of their ZIP code, gender, and date of birth. A table of data, where each column contains a specific type of information (e.g. ZIP code or gender), and each row contains information on a specific individual, is said

to have k-anonymity with respect to a set of quasi-identifiers, if for every combination of quasi-identifiers that appear in the table appear in at least k rows. To illustrate the concept with an example, table 3.1 shows a naively anonymized table while 3.2 shows a k-anonymized table with $k = 2$. The fields {*ZIP code, Gender, Date of Birth*} are quasi-identifiers, while {*Condition*} is a *sensitive attribute*, a piece of information that we don't want revealed about individuals.

The concept of k-anonymity has been extended to location privacy [55]. In the location privacy setting, the location of the user at a given point in time is a quasi-identifier. In effect, a user has k-anonymity if their location trace has is indistinguishable from the location traces of at least $k - 1$ other users. There are a number of problems with k-anonymity both in the context of database privacy and location privacy, which the complementary notions of ℓ-diversity [83] and t-closeness [76] have tried to address. We will not dwell on them here, because k-anonymity has a more fundamental drawback for our purposes. It assumes that the objective of the privacy mechanism is to protect the identity of the user instead of the sensitive attributes themselves.

When protecting a person's identity truly is the objective, k-anonymity can be a helpful metric, but we would argue that for modern location-based services, it is quite likely that the LBS provider already knows the identity of the user and we want to limit what can be inferred about them besides their identity.

Table 3.1: A naively anonymized table where only explicit identifiers such as names and social security numbers have been removed.

ZIP Code	**Gender**	**Date of birth**	**Condition**
13053	M	06/07/1972	Asthma
13068	F	02/11/1972	Diabets
14850	F	04/21/1977	HIV
14867	M	09/13/1977	Cancer

3.3.2 Bayesian Decision Theory

One of the more enduring approaches to quantifying location privacy is based on Bayesian decision theory. Here, we assume that the user's actual location is $x \in \mathcal{X}$, where $\mathcal{X}$ is the space that the user moves around in. It is usually a subset of either $\mathbb{R}^2$ or $\mathbb{Z}^2$. The user doesn't report their true location. The user reports a location $y \in \mathcal{Y}$, in the space of all possible reported locations. In many cases, $\mathcal{Y} = \mathcal{X}$ but that doesn't have to be the case. The mapping from $\mathcal{X}$ to $\mathcal{Y}$ is determined by the privacy mechanism $K : \mathcal{X} \to \mathcal{Y}$. Often the privacy mechanism is stochastic. In that case, the mapping is determined by a conditional distribution $f_K(y|x)$ (in the continuous case) or $p_K(y|x)$ (in the discrete). While $\mathcal{X}$ and $\mathcal{Y}$ can each be continuous or discrete, we will assume for the remainder of the section that both are discrete. All expressions generalize in a straight-forward manner to the continuous case.

As this is a Bayesian approach, we furthermore assume that the adversary has some form of prior π over $\mathcal{X}$. It is theoretically possible to make the prior arbitrarily complex. You can condition it on the time of day/week/month/year, the user's previous location reports, other users' location reports, the weather, and the list goes on. This doesn't change the rest of the theory, but it makes the notation considerably more complex. We will therefore assume an uncondi-

Table 3.2: A k-anonymous table with respect to the quasi-identifiers ZIP Code, Gender, and Date of Birth with $k = 2$. The granularity of the data was reduced until the desired k-anonymity was reached.

ZIP code	Gender	Date of birth	Condition
130**	*	*/*/1972	Asthma
130**	*	*/*/1972	Diabets
148**	*	*/*/1977	HIV
148**	*	*/*/1977	Cancer

tional prior that is purely a function of $\mathcal{X}$.

The consequence of this assumption is that the adversary's inference will depend only on the reported location $y \in \mathcal{Y}$. The inference itself, $\hat{X} = \delta(Y)$, will depend on the chosen *loss function*, $L : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$. Typical loss functions are the 0-1 loss function:

$$L_{0,1}(x, x') = \begin{cases} 0 & x = x' \\ 1 & x \neq x' \end{cases}$$

and the Euclidean distance squared or Mean Squared Error (MSE):

$$L_{MSE}(x, x') = ||x - x'||^2$$

In an influential paper, Shokri et al. [115] argued that the *posterior expected loss* should be the measure of privacy:

$$E[L(X, \delta(Y))|Y = y] = \sum_{x \in \mathcal{X}} p(x|y) L(x, \delta(y))$$

This is of course a function of the reported location y, but the *Bayes risk* is not. We get the Bayes risk by taking the unconditional expectation of the loss function:

$$\begin{aligned} E[L(X, \delta(Y))] &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} L(x, \delta(y)) p_{X,Y}(x, y) \\ &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} L(x, \delta(y)) p_K(y|x) \pi(x) \end{aligned}$$

as described in [116]. If the loss function is the 0-1 loss, we will often use the *expected success rate* (ESR) as a metric instead of the loss function. The ESR is the probability of the correct location being inferred and is the complement of the Bayes risk:

$$\begin{aligned} P(X = \hat{X}) &= 1 - E[\mathbf{1}\{X \neq \hat{X}\}] \\ &= E[\mathbf{1}\{X = \hat{X}\}] \\ &= E[\mathbf{1}\{X = \delta(Y)\}] \\ &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} E[\mathbf{1}\{X = \delta(Y)\} | X = x, Y = y] p_{X,Y}(x, y) \\ &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} L_{0,1}(x, \delta(y)) p_{X,Y}(x, y) \\ &= E[L(X, \delta(Y))] \end{aligned}$$

The Bayes-optimal inference, $\delta : \mathcal{Y} \to \mathcal{X}$, depends on the loss function. For the 0-1 loss function, it can be shown that the maximum a posteriori (MAP) inference is optimal:

$$\begin{aligned} \delta_{0,1}(y) &= \arg\max_{x \in \mathcal{X}} p_{X|Y}(x|y) \\ &= \arg\max_{x \in \mathcal{X}} \frac{p_{Y|X}(y|x)\pi(x)}{\sum_{x' \in \mathcal{X}} p_{Y|X}(y|x')\pi(x')} \\ &= \arg\max_{x \in \mathcal{X}} p_{Y|X}(y|x)\pi(x) \end{aligned}$$

For the MSE loss, it can be shown that the mean of the posterior expectation gives the optimal inference:

$$\begin{aligned}\delta_{MSE}(y) &= E[X|Y=y] \\ &= \sum_{x\in\mathcal{X}} x p_{X|Y}(x|y) \\ &= \frac{1}{\sum_{x'\in\mathcal{X}} p_{Y|X}(y|x')\pi(x')} \sum_{x\in\mathcal{X}} x p_{Y|X}(y|x)\pi(x)\end{aligned}$$

The Bayesian approach has a lot of advantages. It is a well-known theory, is quite intuitive and easy to interpret, and it makes it easy to quantify our assumptions about the adversary's goals and prior knowledge. Furthermore, it usually results in tractable expressions that can easily be computed with the plethora of computational tools developed for Bayesian inference. However, it does require us to make explicit assumptions about the adversary's prior knowledge. In the next section, we will see a different approach that obviates the need to make such assumptions.

3.3.3 Differential privacy

Differential privacy has become the de facto gold standard for privacy. It was introduced in 2006 by Cynthia Dwork as a way to answer statistical queries from databases in a manner that protected the privacy of the individuals in the database [38]. Similarly to k-anonymity, a database is a table with a number of rows and columns. Each row contains information about a single individual and each column contains information about a single attribute. The idea behind differential privacy is to answer statistical queries about the database in a way that is virtually unaffected by the addition or removal of a single row. We say a privacy mechanism K provides ϵ-differential privacy if for any databases D_1, D_2

that differ by at most one row, and for any $S \subseteq \text{range}(K)$,

$$P(K(D_1) \in S) \leq e^{\epsilon} P(K(D_2) \in S)$$

By symmetry, this implies that

$$e^{-\epsilon} \leq \frac{P(K(D_1) \in S)}{P(K(D_2) \in S)} \leq e^{\epsilon}$$

Informally, K provides differential privacy if the probability of any particular output is very similar for any two databases differing only by one row. The number of rows that two databases differ by is called the Hamming distance within the field of databases and is sometimes denoted as $d_h(\cdot, \cdot)$. Differential privacy can thus be stated slightly more generally as follows: For any two databases D_1, D_2, and any $S \subseteq \text{range}(K)$,

$$P(K(D_1) \in S) \leq e^{\epsilon d_h(D_1, D_2)} P(K(D_2) \in S)$$

As a concrete example, differential privacy can be used to guarantee that the published outcome of a survey will not differ substantially based on any one individual's participation, so they might as well participate.

Adding Laplacian noise to data is a common method for obtaining differential privacy [39]. It goes under the name *Laplace mechanism*. To define the Laplace mechanism, we first must define the concept of a query function. A query function is a function $f : \mathcal{D} \to \mathbb{R}$, where $\mathcal{D}$ is the set of admissible databases. For example, given a database of names, ages, and blood pressure, a

query function f could answer the question "what is the average systolic pressure of people over the age of 50". The *sensitivity* Δf of a query function f is defined as the maximum absolute difference in response between two databases D_1, D_2, differing by at most one row:

$$\Delta f := \max_{\substack{D_1, D_2 \in \mathcal{D} \\ d_h(D_1, D_2) \leq 1}} ||f(D_1) - f(D_2)||$$

Laplacian noise is a zero-mean Laplacian random variable with scale parameter σ and a PDF given by

$$f_{Lap}(x|\sigma) = \frac{1}{2\sigma} \exp\left(-\frac{|x|}{\sigma}\right)$$

Adding Laplacian noise with scale parameter $\sigma = \Delta f/\epsilon$ provides ϵ-differential privacy. Let D_1, D_2 be arbitrary databases in $\mathcal{D}$ such that $d_h(D_1, D_2) \leq 1$. And let f be a query function with sensitivity Δf and define $x_1 := f(D_1)$ and $x_2 := f(D_2)$. Furthermore, let p_i be the probability density function (PDF) of $K(f(D_i)) = x_i + W_i$ for $i = 1, 2$ where W_1, W_2 are independent and identically distributed (i.i.d.) zero-mean Laplace random variables with scale parameter $\sigma = \Delta f/\epsilon$ for some $\epsilon > 0$. Then for any $y \in \mathbb{R}$, we have

$$
\begin{aligned}
\frac{p_1(y)}{p_2(y)} &= \frac{\exp\left(-\frac{\epsilon}{\Delta f}|x_1 - y|\right)}{\exp\left(-\frac{\epsilon}{\Delta f}|x_2 - y|\right)} \\
&= \exp\left(\frac{\epsilon}{\Delta f}(|x_2 - y| - |x_1 - y|)\right) \\
&\leq \exp\left(\frac{\epsilon}{\Delta f}|x_1 - x_2|\right) \\
&\leq \exp\left(\frac{\epsilon}{\Delta f}\Delta f\right) \\
&= \exp(\epsilon)
\end{aligned}
$$

Local Differential Privacy

An implicit assumption behind differential privacy is that users give their data to a trusted third party that aggregates and anonymizes the data prior to releasing it to the adversary. If the LBS is the adversary, that is not going to work. In our setting, users are releasing their data directly to the adversary, one data point at a time. The original form of differential privacy as described above is poorly suited to such a setting. Instead, a stronger notion of differential privacy has been formulated for such a setting called *local differential privacy* [cite]. In this case, the privacy mechanism K is applied to individual data points in $\mathcal{X}$, the domain of K, to produce a point in $\mathcal{Y}$, the range of K. A privacy mechanism K is said to be ϵ-locally differentially private if for all $x_1, x_2 \in \mathcal{X}$ and all $y \in \mathcal{Y}$,

$$P(K(x_1) = y) \leq e^{\epsilon} P(K(x_2) = y)$$

The classical method of obtaining local differential privacy is via the *randomized response* mechanism [128]. The randomized response mechanism works

when x_1, x_2 are binary attributes, i.e. $\mathcal{X} = \{0, 1\}$. The idea is that with probability p, the user answers truthfully and with probability $1 - p$, they report the wrong answer. In other words, $P(K(x) = x) = p$ and $P(K(x) \neq x) = 1 - p$. This method was originally proposed in the context of making surveys on sensitive or potentially illegal topics, e.g. "Have you ever smoked marijuana?".

To see that the randomized response mechanism provides local differential privacy, let $p \geq 0.5$. Then for any $x_1, x_2 \in \mathcal{X} = \{0, 1\}$ and any $y \in \mathcal{Y} = \mathcal{X}$, we have

$$\max_{x_1, x_2, y \in \mathcal{X}} \frac{P(K(x_1) = y)}{P(K(x_2) = y)} = \frac{p}{1 - p}$$

Therefore K provides ϵ-local differential privacy for $\epsilon = \log \frac{p}{1-p}$. If $p < 0.5$, then $\epsilon = \log \frac{1-p}{p}$.

Geo-indistinguishability

Differential privacy tends to prescribe mechanisms that perturb the data very heavily, to the degree that in many cases the result is utterly useless [8]. Because of this, a number of relaxations of differential privacy have been developed [39, 17, 90]. One such relaxation, with the jaw-breaking name *geo-indistinguishability*, is specifically formulated for location privacy [4]. It can be applied when the Euclidean metric is well-defined on $\mathcal{X}$. A privacy mechanism K is geo-indistinguishable if for all $x_1, x_2 \in \mathcal{X}$ and all $S \subseteq \mathcal{Y}$,

$$P(K(x_1) \in S) \leq e^{\epsilon d(x_1, x_2)} P(K(x_2) \in S)$$

where $d(\cdot, \cdot)$ is the Euclidean metric. In $\mathbb{R}^2$, this can be achieved by adding planar Laplacian noise with scale parameter $\frac{1}{\epsilon}$. The planar Laplace distribution has the following density in polar coordinates [4]:

$$f(r, \theta) = \frac{r}{2\pi\sigma^2} \exp\left(-\frac{r}{\sigma}\right) \tag{3.1}$$

Andrs et al. propose that ϵ is determined by deciding on a radius R within which we want ϵ'-differential privacy. Then set $\epsilon = \epsilon'/R$. This leads to an alternate way to characterize geo-indistinguishability, one that mirrors local differential privacy just as the above formulation mirrors classical differential privacy. If the range of K, $\mathcal{Y}$, is equal to $\mathcal{X}$, then for any $x_1 \in \mathcal{X}$ and any $x_2 \in B(x_1, R)$, the ball of radius R centered on x_1, and any $y \in \mathcal{Y}$,

$$f_K(y|x_1) \leq e^{\epsilon'} f_K(y|x_2)$$

3.3.4 Quality of Service

In this section we cover a few common methods of measuring quality of service.

(α, δ)-usefulness

When the location privacy mechanism is stochastic rather than deterministic, it makes sense to have a probabilistic measure of quality of service. When adding Gaussian or Laplacian noise, for example, there is theoretically no absolute upper bound on the distance between the reported location and the ac-

tual location. In those cases we can adopt the notion of (α, δ)-usefulness, introduced in [15]. We say a mechanism K is (α, δ)-useful if the perturbed location is within a distance of α of the real location with probability at least δ. More formally, let $K : \mathcal{X} \times \Omega \to \mathcal{X}$ be a privacy mechanism. Then we say K is (α, δ)-useful if $P(d(K(x), x) \leq \alpha) \geq \delta$. A mechanism thus preserves more utility as it approaches $(0, 1)$-usefulness, while the QoS decreases as it approaches $(\infty, 0)$-usefulness.

Expected QoS degradation

The other approach is to measure the QoS degradation as the expected distance between the true location and the reported location under some metric. Assuming $\mathcal{X} = \mathcal{Y}$, then we define the expected QoS degradation as

$$Q(K, \pi) = E[d(X, Y)] = \sum_{x,y \in \mathcal{X}} d(x, y) p_{X,Y}(x, y)$$

where d is some metric on X. Reasonable choices include the Euclidean metric and the Euclidean metric squared.

3.4 Semantic Location Privacy

Protecting physical location privacy is a great first step. It is a natural starting point for location privacy and its relatively easy to reason about and analyze. We have a plethora of tools to analyze the geometry of $\mathbb{R}^2$ and $\mathbb{Z}^2$, so most analyses result in easily computable expressions. However, as we argued in the introduction, the physical location is not necessarily what the user wants or needs to protect. Using the expected physical distance as a metric for privacy may be only indirectly related to the actual sensitive information the user needs to protect. The expected distance between the reported location and the actual location could be 1 km, but that would provide very different levels of privacy in a densely populated urban area and a sparsely populated rural area.

This is why we take a different approach in this dissertation. We make the semantics of a user's location the center of our analysis and design of location privacy mechanisms. Thus, instead of focusing on physical location (e.g. latitude 42.44064, longitude -76.49661) , whether continuous or discretized, we focus on protecting the information about which actual places the user visits, e.g. "Starbucks on Main Street, Springfield, USA" or "Rhodes Hall, Cornell University, Ithaca, NY". We call this the *semantic location* of a user.

In order to make semantic locations easier to work with, we associate a physical location with each semantic location. In the rest of this chapter, a set of semantic locations will refer to the set of physical locations associated with the semantic locations. One might argue that a semantic location such as a coffee shop or a supermarket is not just associated with a single physical location, but an infinite number covering the area of the semantic location. While it is pos-

sible to model the physical location associated with a semantic location as a set of coordinates, or even more generally, a probability distribution over a set of coordinates, I believe that the added complexity introduced by such a faithful model outweighs any benefits conferred by it. we will therefore stick to associating semantic locations with a single physical location and will refer to it as the *point-mass assumption*.

The set of semantic locations in a given area can be ambiguous, however. Is a mall a semantic location, or is each individual shop inside the mall a semantic location? When does an area become a semantic location? It is inherently difficult to give a mathematically precise answer to these questions. We can, however, turn to crowdsourcing. Location-based social media such as Foursquare[1], Yelp[2] and Facebook Places[3] have harnessed crowdsourcing to create a consensus about what constitutes a semantic location. Even better, they provide us with an indication of the relative popularity of each semantic location which can be used to construct an aggregate prior.

Given this new approach to location privacy, we want to measure, analyze, and mitigate semantic location privacy issues. We need to adapt the location privacy and quality of service metrics of physical location privacy to semantic location privacy. And we want to develop mechanisms optimized for semantic location privacy based on our adapted metrics. In section 3.5 we present an adaptation of the Bayesian approach and use data collected from location-based social networks to create $\mathcal{X}$ and π. In section 3.6 we present a theoretical framework based on stochastic geometry, where we model $\mathcal{X}$ and π with point processes. This approach doesn't require crowdsourced data to construct the

[1] www.foursquare.com
[2] www.yelp.com
[3] www.facebook.com/places

set of semantic locations and provides insights into designing semantic location privacy mechanisms that the empirical approach does not. In section 3.7 we describe how to adapt differential privacy to semantic location privacy and in section 3.8 we describe a location privacy mechanism specifically designed for semantic location privacy.

3.4.1 Notation

The set of semantic locations, $\mathcal{S}$ is a finite or countable subset of $\mathcal{X}$, which in turn is a subset of $\mathbb{R}^2$. The adversary's prior π, is now over $\mathcal{S}$ instead of all of $\mathcal{X}$. Location privacy mechanisms are denoted by K, and map from $\mathcal{S}$ to the set of possible reported locations $\mathcal{Y}$, which would typically be equal to $\mathcal{S}$ or $\mathcal{X}$. We will assume that the prior π is an accurate measure of the probability of a location report from an arbitrary user at an arbitrary time comes from location $s \in \mathcal{S}$. The adversary's goal is now to uncover the user's true location $s \in \mathcal{S}$ given the reported location $y \in \mathcal{Y}$. We assume the adversary either already knows the identity of the user or simply does not care.

3.4.2 Measuring QoSD

We will use a slight modification of the (α, δ)-usefulness metric as our preferred QoSD metric for semantic location privacy, which we call the *area of uncertainty*. For deterministic mechanisms like quantization, the area of uncertainty will simply be the size of the quantization bins:

$$\text{QoSD}(\Delta) = \Delta^2$$

For probabilistic mechanisms, the area of uncertainty is the smallest area that contains the user with 90% probability:

$$\text{QoSD}(\sigma) = \pi\alpha^2$$

where α is such that $P(||W|| \leq \alpha) \geq 0.9$.

3.4.3 Related Works

In [19], the authors try to infer semantic locations, i.e. meaningful locations, from raw GPS data from a large set of users. The papers [61] and [79] similarly try to extract semantic locations from physical location data (GPS) from a set of users. The paper [24] proposes a similar semantic location framework, but its analysis and mechanism design is focused on protecting user anonymity. The concept of ℓ-diversity as it is applied to location-privacy [137] is also closely related to semantic location privacy. It is a metric that is meant to complement k-anonymity, by ensuring that a group k-anonymous records has at least ℓ distinct values of the sensitive attribute(s). However, its goal is to ensure better anonymization of location data, not protecting semantic information content. In [75], the authors propose to measure semantic location privacy as the Earth Mover's Distance (EMD) between the adversary's prior and posterior and describe a privacy mechanism that implements a non-uniform quantization scheme that ensures the EMD of the adversary will not increase more than

a set threshold. The problem with EMD is of course that it only captures the distance between two probability distributions, not whether the posterior gives a ”better” picture of where the user has been. The idea of p-sensitivity is another extension of k-anonymity that takes semantic locations into account, but is also focused on protecting anonymity of users [136]. The PROBE approach, described in [30] describes a scheme that allows a user to specify certain areas as ”sensitive” and makes it harder for an adversary to infer when the user is within a sensitive area.

3.5 Empirical Bayesian Analysis

In this section, we evaluate location privacy mechanisms with a Bayesian approach using empirical data gathered from location-based social media. This work has previously been published in the proceedings of the 49th Hawaii International Conference on System Sciences [67].

Location-based services provide information such as directions, recommendations, and weather based on the user's current position. In the process of providing this information, the service can accumulate a detailed location trace of the user. Such a trace can potentially reveal a host of personal information about the user even if the user never intended to share that information. It is therefore of the utmost importance to take location privacy into account when providing or using such services.

The past decade has seen a lot of important work on location privacy. Proposed schemes include reducing the granularity of the location data [55], adding noise [4], adding position dummies [69], all of which obfuscate the user's location. However, much less work has focused on *how much* obfuscation is needed to prevent specific inference attacks. The work that has been done in that area focuses mostly on k-anonymity [123], much like the schemes mentioned above. While providing k-anonymity can be valuable, it is often neither sufficient [117] nor relevant: in many cases, the users wish to register their own identity with the location-based service, for example in location-based social networks. In those cases, k-anonymity is of limited use. However, the users might still want to limit the scope and scale of personal information that can be gleaned from their location traces. In this chapter, we focus on limiting semantic location

attacks. A semantic location is a location that holds some sort of meaning to users, such as Starbucks on Main St. or Central Park. This is opposed to physical location, typically in the form of WGS 84 coordinates such as latitude 40.782, longitude 73.965.

It stands to reason that most types of obfuscation will suffice to resist simple semantic location attacks such as reverse geocoding. However, the adversary can use map information to diminish the effect of the location obfuscation by excluding inaccessible areas from consideration [130] [31]. In fact, the adversary can go a step further and determine all the semantic locations in the area. The user is of course not equally likely to be at all semantic locations, so a probability distribution over them must be determined. We construct such a probability distribution using Foursquare venue data from five U.S. cities of varying sizes and examine the posterior probability of Bayesian location inferences for different location resolutions. Our results indicate that location must be heavily quantized to guarantee a low posterior probability in location inferences, but if one is willing to settle for a low *expected* posterior probability, the resolution can be increased by an order of magnitude. Furthermore, we show that an adversary with a prior distribution over semantic locations can expect to do significantly better than one who simply assumes a uniform prior. The rest of the chapter is organized as follows. In section 3.5.1, we cover related works. Our assumptions about the adversary are made explicit in section 3.5.2. Section 3.5.3 describes the dataset used, section 3.5.4 the experiments performed, and section 3.5.5 the main results. Section 3.5.6 concludes the chapter.

3.5.1 Related Work

Previous studies of the effect of location granularity on inference attacks have mostly focused on inferring the identity of the user, as opposed to the semantic locations visited. The work closest to ours examines how much obfuscation is necessary to prevent an adversary from inferring a user's home address based on car GPS traces [72]. The effect of location granularity on the k-anonymity of home/work location pairs have also been examined using anonymized data from the U.S. Census bureau [54], and on the k-anonymity of the user's top n most visited locations using call detail records [141]. In both studies, it was found that location cannot be more precise than city or county level before compromising k-anonymity.

Location privacy-providing mechanisms form a distinct, yet related class of works. In these works, the goal is usually to provide a measure of location privacy as measured by a metric such as k-anonymity or expected distance error. There exist excellent review papers for general location privacy-providing mechanisms [130] [73], we mention merely the ones most related to our study. A number of mechanisms extend the concept of l-diversity [83] to location privacy by requiring at least l distinct locations in the obfuscated area [137] [7] or l categories of semantic locations, such as schools, churches, offices, etcetera [75] [24]. Semantic location inferences have also been studied in non-adversarial settings, i.e. assuming the user wants the location inferences to succeed. Some results indicate that at least the category of semantic location can be inferred with a high success rate by drawing on all available smartphone data, including personal calendar and Internet browsing history [61] [142]. We do not include such side information in our study, but it might be the topic of a future study.

(a) Sedona, AZ

(b) Ithaca, NY

(c) Syracuse, NY

(d) St. Louis, MO

(e) Manhattan, NY

Figure 3.2: Streetmaps on a 1:50,000 scale. Source: [46]

3.5.2 Threat Model

This section lays out our assumptions about the adversary. The adversarys goal is to infer all non-residential semantic locations the user has been to. We limit our scope to non-residential locations and leave our users homes or the homes of their friends. The reasons are twofold. Inferring users home addresses has been treated in previous work [72] [79]. Furthermore, using a prior location distribution would not necessarily be helpful in inferring residential locations, as the distribution over them could be very flat. The goal behind residential semantic locations would undoubtedly be to infer associations between people. It can be considered as a separate problem requiring separate tools to carry out and prevent. Thus the adversary focuses solely on non-residential locations.

To this end, the stay locations have already been extracted from the location trace, for example using thresholding [79] [143] or clustering [19] [142]. The home location can readily be identified and removed from the list of stay locations [72], while other homes visited will be counted as noise in the inference.

The location coordinates reported by the user have obfuscated by quantization as described further in section 5. The adversary uses a prior distribution over all possible non-residential semantic locations to infer the most likely semantic location using the Maximum A Posteriori (MAP) rule. The adversary does not have any other information about the user beside the location trace. Our goal is to limit the confidence with which the adversary can make the inference, as measured by the posterior probability of the most likely semantic location.

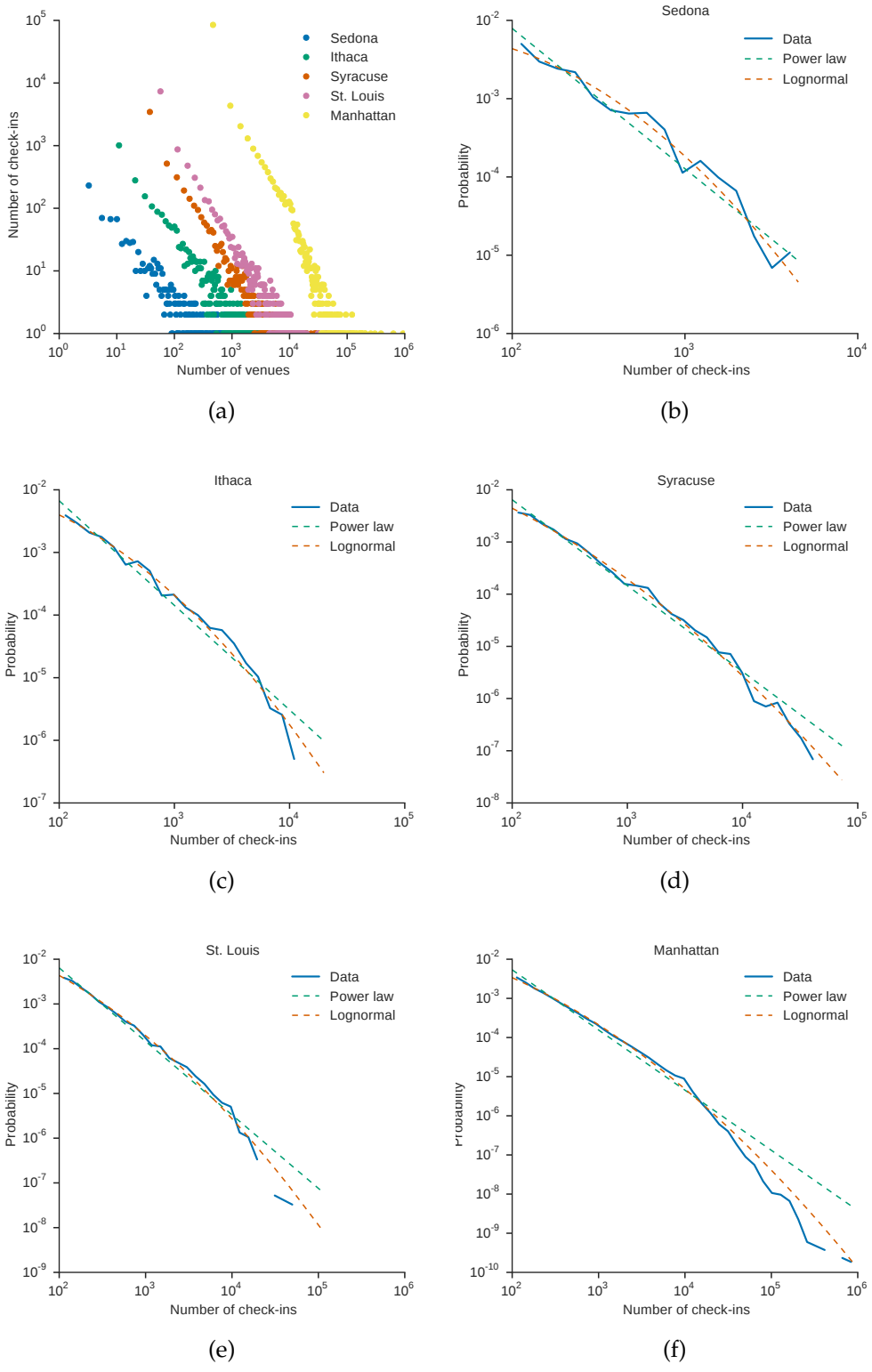


Figure 3.3: The distribution of check-ins per venue. (a): All cities on a double-logarithmic scale. (b)-(f): Empirical and fitted probability density functions on double-logarithmic scales.

3.5.3 The Dataset

We used Foursquare [47] venue data to build the dataset of semantic locations, as Foursquare has one of the most complete databases of semantic locations and includes and indicator of the locations' relative popularity in the form of "check-in" data.

Foursquare users check in to the semantic locations they are at, allowing them to see who else is there, advertise their location to their friends, and earn points and discounts. The service has over 45 million registered users [48] and is one of the most popular location-based social networks [144]. Its dataset is frequently used in academic studies, including research on urban planning [25] [108] [97] [16], factors affecting venue popularity [81] [77], and to infer urban activity [97].

The locations users can check in to are called venues in the Foursquare system. Each venue has a number of attributes associated with it, including name, unique ID, category, latitude and longitude, user count, and total check-ins. We use the total number of check-ins to construct the prior probability distribution over the venues (i.e. the semantic locations).

Admittedly, the prior is biased in at least two ways. First, the Foursquare demographic is not necessarily the same as that of the general population. It is reportedly biased slightly towards males; people 18-29 years of age; and the college-educated [50]. However, the demographics of general location-based service users is similarly biased [144]. Secondly, since all check-ins are voluntary, there might be a self-report bias present. It is easy to imagine that a user is more likely to check into a popular restaurant than a therapists office. Despite

Table 3.3: A list of cities in the study along with their population, the coordinates of the bounding box of the 10 x 10 km area centered on the city from where the data was gathered, and the total number of semantic locations (venues) in that area. Population data source: U.S. Census Bureau [22]

City	Pop.	S. lat.	N. lat.	W. lon.	E. lon.	Venues
Sedona	10,031	34.8094	34.8994	-111.8504	-111.7404	991
Ithaca	30,014	42.3994	42.4894	-76.5537	-76.4317	2,690
Syracuse	145,170	43.0031	43.0931	-76.2074	-76.0844	5,734
St. Louis	319,294	38.5820	38.6720	-90.2950	-90.1790	10,704
Manhattan	1,585,873	40.6960	40.7860	-74.0509	-73.9309	98,662

these biases, Foursquare is still arguably the best publicly available dataset on the relative popularity of semantic locations and can be used as long as one is aware of its limitations.

Our dataset consists of all Foursquare venue data in 10-by-10 kilometer areas centered on five U.S. cities of varying sizes: Sedona, AZ; Ithaca, NY; Syracuse, NY; St. Louis, MO; and Manhattan, NY. Table 3.3 shows the geographical coordinates bounding each area, as well as city population and number of venues gathered. The streetmaps of the areas in question are shown in figure 3.2 and figure 3.3 shows the distribution of total check-ins on a log-log scale. Previous studies have found the distribution of check-ins to be well-modeled by power laws or lognormal distributions [77] [129]. Which distribution is more appropriate is a question of on-going debate [92] [129].

3.5.4 Experiments

Using the Foursquare dataset described above, we conduct an experiment to examine the effect of location granularity on the confidence with which an adversary can carry out semantic location inferences. Each area is divided into

$n \times n$ square cells for varying numbers of n and it is assumed that the user only reports the cell they are located within, for example by reporting the coordinates of the center of the cell. The grid size, i.e. the size of the cell as measured by its side length in meters, becomes our measure for the location granularity.

For each cell, we look at the conditional probability of the most likely semantic location within the cell, i.e. the maximum a posteriori probability. We then compute the maximum, mean, median, and expected MAP over all the cells as a function of grid size (i.e. location granularity).

Before we go on, let us explain our mathematical notation. Let $\mathcal{X}$ be the area in question, which can be thought of as a solid square in $\mathbb{R}^2$. The semantic locations $\mathcal{S} = \{s_1, ..., s_m\}$ are points within $\mathcal{X}$. We have a prior user location distribution, π, over $\mathcal{X}$, with support only on $\mathcal{S}$. We conduct the experiments using both a uniform prior π_u, where every point in $\mathcal{S}$ is equally likely, and using the Foursquare prior, π_f, where the probability of each point in s is proportional to the number of user check-ins. Specifically, if checkins(s_k) is the number of check-ins for location s_k, then

$$\pi_f(s_k) = \frac{\text{checkins}(s_k)}{\sum_{i=1}^{m} \text{checkins}(s_i)}$$

for the Foursquare prior and

$$\pi_u(s_k) = \frac{1}{|\mathcal{S}|}$$

for the uniform prior, and zero everywhere else in $\mathcal{X}$. Here, $|\mathcal{S}|$ denotes the number of elements in $\mathcal{S}$ (the cardinality). The quantization happens by partitioning $\mathcal{X}$ into $n \times n$ cells, $C_1, ..., C_{n^2}$, for increasingly large n. The MAP rule used by the adversary can be expressed as follows:

$$\hat{s} = \arg\max_{s \in C} \frac{P(s)}{\sum_{s' \in C} P(s')}$$

where C is the cell reported by the user. The posterior probability of the most likely semantic location inference is then simply a function of the cell as computed by

$$P_{MAP}(C) = P(\hat{s}|C).$$

We run the experiments using both π_u and π_f in place of $P(\cdot)$. For the simple case of a uniform prior, the posterior probability is just the reciprocal of the number of semantic locations within C. When using the Foursquare prior, the relative popularity of the semantic location influences the estimate towards the more popular locations.

We go on to compute P_{MAP} for every cell containing at least one semantic location. We then compute the minimum, median, mean, and expected P_{MAP} over all cells with at least one semantic location. The expected error probability is computed by a weighted average over all cells, where the weight is given by the probability of a user being within the cell. In other words, we compute the Expected Success Rate (ESR) as described in 3.3.2:

$$P(X = \hat{X}) = E[P_{MAP}(C)] \tag{3.2}$$

$$= \sum_{i=1}^{n^2} P_{MAP}(C_i) P(C_i) \tag{3.3}$$

These experiments are performed for all five cities. An overview of our results can be seen in figure 3.4.

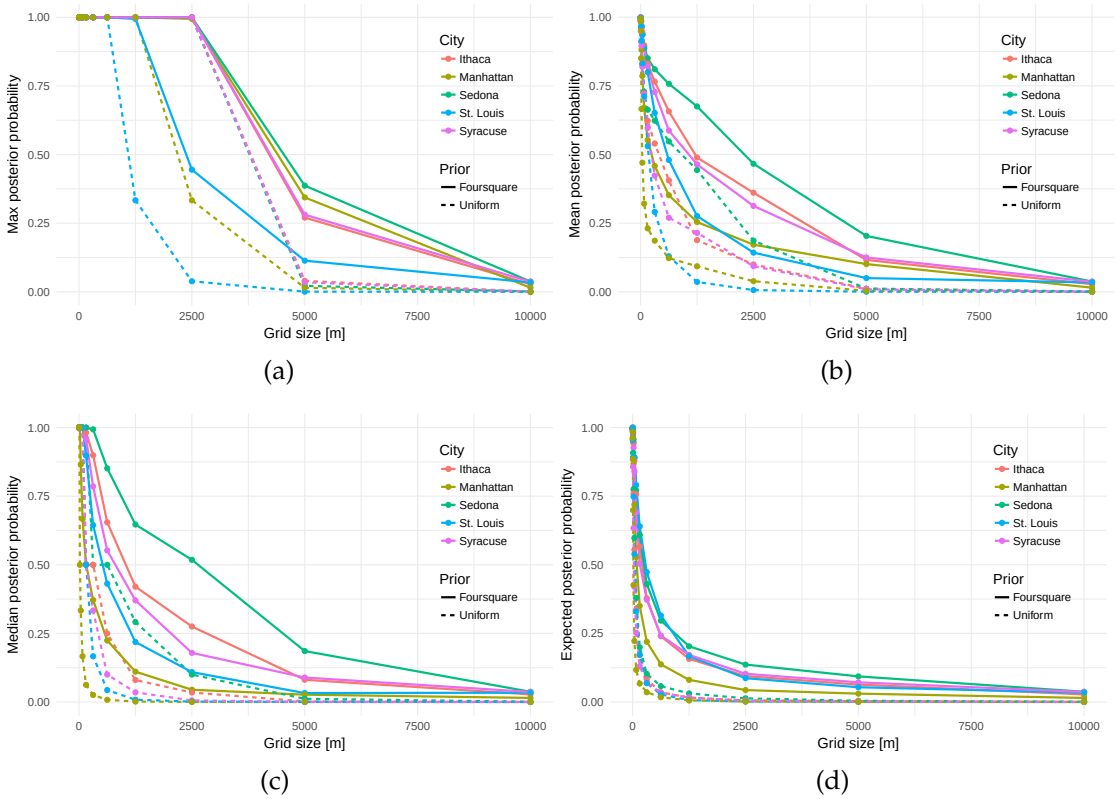


Figure 3.4: The minimum (a), mean (b), median (c), and expected (d) posterior probability as a function of grid size. The grid size must be on the order of 5 km to achieve a significant drop in the maximum confidence an adversary can have in the semantic location inference. The median posterior probability drops faster, and the expected probability (as computed in eqn. 3.3) drops faster still.

3.5.5 Results

The evaluation of the results poses a certain challenge. We do not have a set of actual, fine-grained location traces from a randomly sampled subset of the population. In fact, the Foursquare venue data represents our best estimate of the actual underlying probability distribution over the semantic locations. Thus, assuming that the Foursquare distribution is accurate, we evaluate the confidence with which the adversary can make semantic location inferences, assuming the adversary knows the underlying distribution.

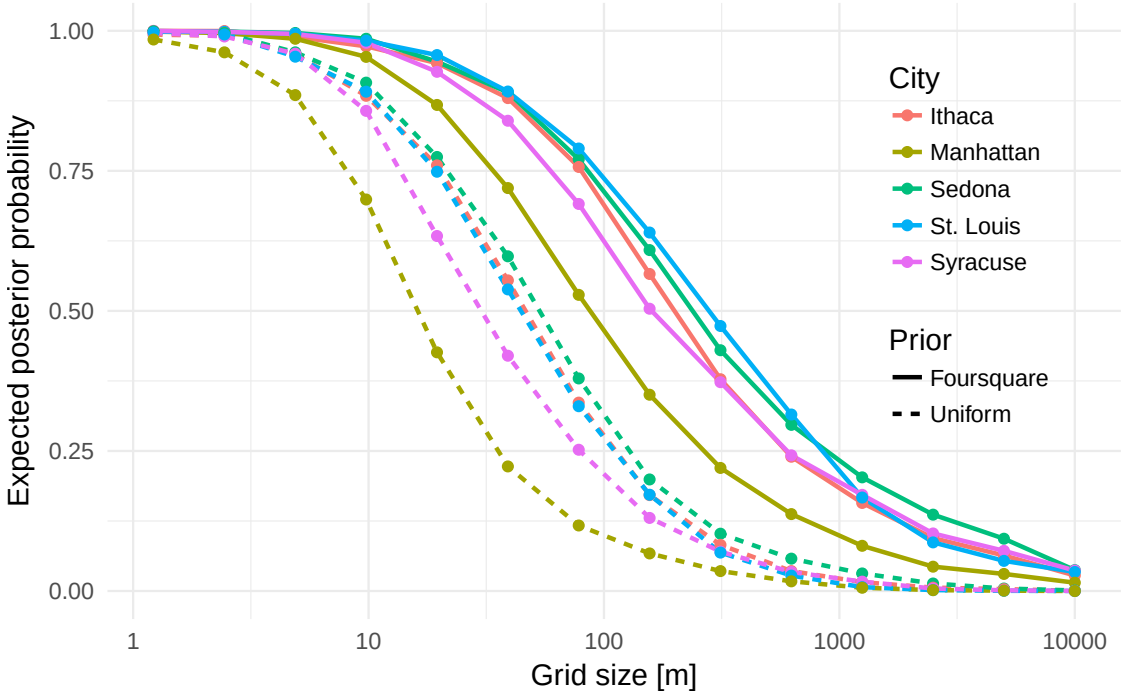


Figure 3.5: The expected a posteriori probability on a semilogarithmic scale. It has already dropped below 0.5 at a grid size of $10^{2.5} \approx 300$ meters for all cities.

Under those assumptions, we see that, unsurprisingly, the adversary's confidence in location inferences varies with city size and location granularity. More interestingly, we see that in order to avoid all location inferences, the location data must be heavily quantized, i.e. the grid size must be very large. As can

be seen in figure 3.4(a), the maximum posterior probability does not drop significantly in most cities until the grid size reaches approximately 5000 meters. This agrees with previous results on inferring home addresses [72], and is of a similar order of magnitude as the necessary granularity to provide a reasonable k-anonymity level [54] [141].

However, the *expected* posterior probability is arguably the more natural measure of the adversary's confidence in the inference. It can be interpreted as the weighted mean posterior probability where more weight is given to locations where the user is more likely to be at. Thus, assuming the users follow the Foursquare distribution, a low expected posterior probability will mean that most of the time, the adversary will have low confidence in the location inferences.

The results indicate that if one is willing to accept a low expected posterior probability, as opposed to low maximum posterior probability, the cell size can be reduced by an order of magnitude. Figures 3.4(d) and 3.5 show that inferences become significantly more difficult at a resolution of around 500 meters. This could be explained by the fact that in the Foursquare model, users are more likely to be in areas with a high density of semantic locations. When the location data is quantized, those semantic locations end up in the same grid, causing a higher uncertainty about the specific semantic location.

As can be expected, the results also show that location inferences are harder in larger cities, and easier using the Foursquare prior than a uniform prior. In Manhattan, the expected posterior probability drops below 20% at 300 meters using the Foursquare prior, while in most other cities it takes a resolution of around 1200 meters to get to a similar level. Similarly, the expected posterior

probability drops below 20% for most cities at a resolution around 300 meters using the uniform prior and around 1500 meters using the Foursquare prior.

3.5.6 Conclusion

We set out to examine how much quantization is necessary to deter semantic location inferences given an adversary with a prior distribution over semantic locations. Assuming that the probability distribution constructed from the Foursquare venue data accurately reflects the aggregate behavior of users of location-based services, we were able to give some estimates of the scale at which the adversary loses confidence in the predictions. The maximum, mean, and median confidence - as measured by the a posteriori probability - dropped quite slowly and required the quantization grid size to be on the order of several kilometers before the confidence dropped below even 0.5. However, if we consider the expected posterior probability, the confidence had dropped below 0.5 in all cities at a grid size of around 500 meters. Thus, given our assumptions, it appears that rounding longitude and latitude to the nearest 500 meters will make most - but not all - semantic location inferences difficult. These results are, to the best our knowledge, the best estimate of the amount of quantization necessary to prevent semantic location inferences.

The 500 meter cutoff stands in contrast to the quantization necessary to provide k-anonymity, which is at the city or county level. It is therefore conceivable that users can report their location with much more precision and still have some degree of privacy, even though their location trace may be unique to them, since the semantic locations will hard to infer. This could be impor-

tant to location-based services wanting to provide privacy without sacrificing quality of service.

There is certainly more work to be done in the area of semantic location inferences. For example, the rich Foursquare dataset could possibly be used to construct privacy-protecting mechanisms that offer semantic location privacy but still allow users to report their location with reasonable accuracy. Furthermore, location privacy schemes other than quantization could be evaluated against an adversary equipped with a Foursquare distribution.

3.5.7 Appendix

Since the publication of the work in this section, I have carried out additional experiments on the collected datasets. Figure 3.6 shows trade-off curves for quantization, additive Gaussian noise, and additive Laplacian noise using the data collected from Ithaca, NY. The Gaussian mechanism has a uniformly better trade-off curve than the Laplace mechanism, which has a uniformly better trade-off curve than quantization.

Figure 3.7 compares the trade-off curves from all five cities using the Gaussian mechanism. Larger cities tend to have better trade-off curves. This makes intuitive sense. Greater density of semantic locations necessitates less perturbation to confuse an adversary.

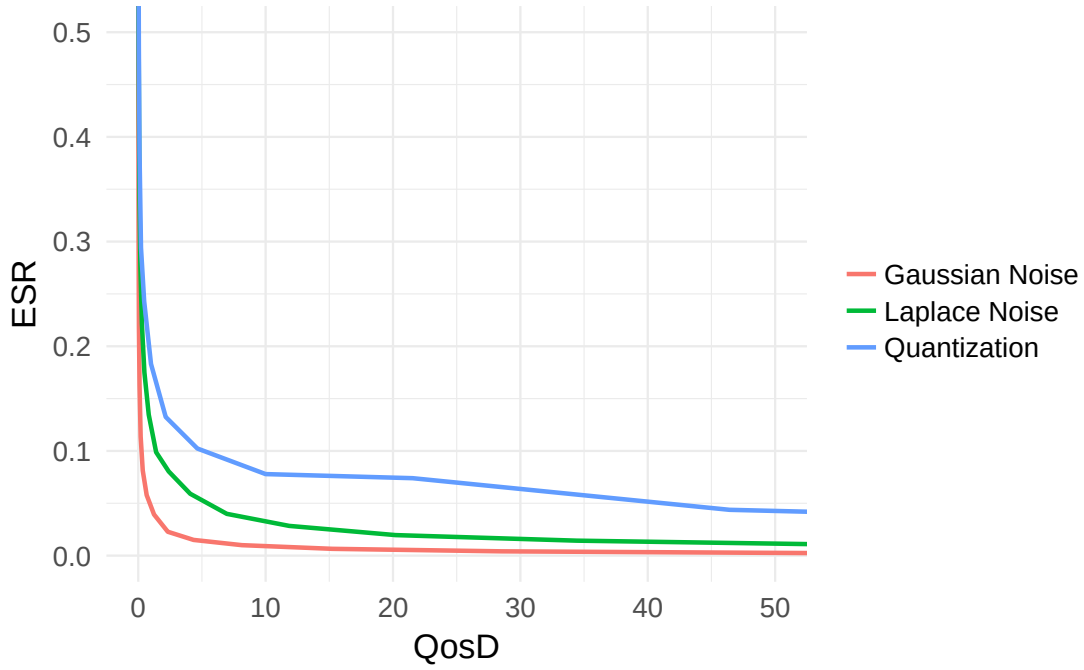


Figure 3.6: Comparison of the privacy-utility trade-off curves for three different mechanisms using data from Ithaca, NY.

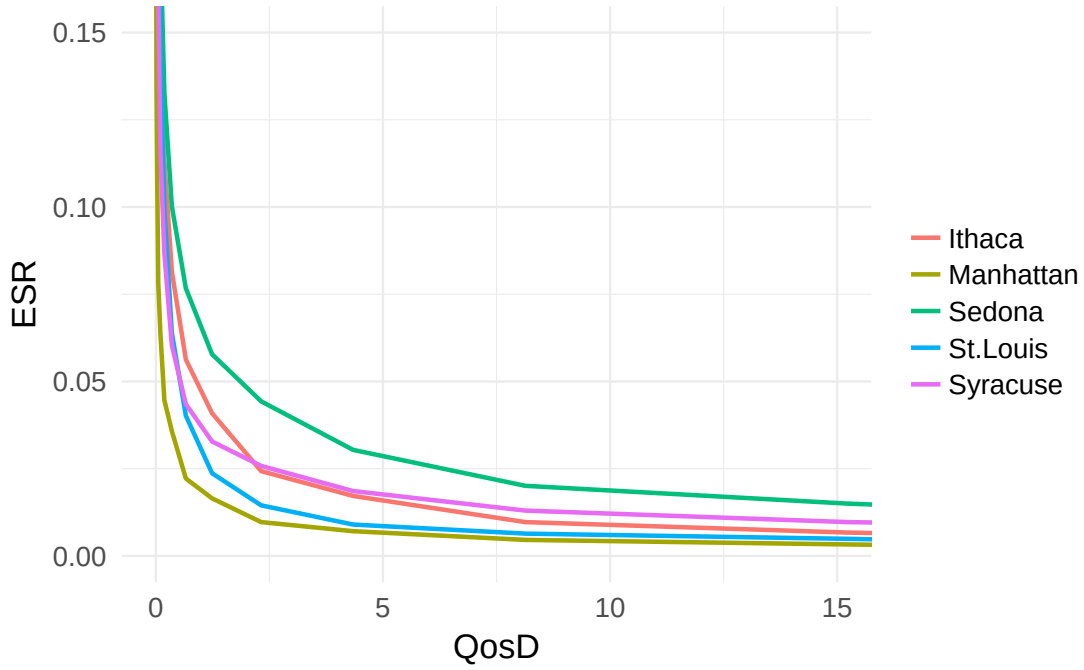


Figure 3.7: Comparison of the privacy-utility trade-off curves of the Gaussian mechanism for all five cities.

3.6 Stochastic Geometry Analysis

Empirical approaches to evaluating semantic location privacy, such as the one described in the previous section, have certain advantages. As described above, data from location-based social networks has certain biases built in, and the end results are dependent on the city or cities we collect data from. Ideally, we have a method of evaluating semantic location privacy mechanisms that is independent from empirical data. This requires us to place location privacy on a firmer theoretical foundation. Such a foundation would help us design, evaluate, and compare obfuscation mechanisms using analytical expressions or simulations instead of relying only on empirical models.

In this section, we present a theoretical framework to analyze semantic location privacy based using stochastic geometry. We model $\mathcal{S}$ as a spatial point process on $\mathcal{X}$ and use the existing theoretical machinery on Poisson point processes to derive insights about our privacy mechanisms. The result is that both $\mathcal{S}$ and the prior over it, π are stochastic. By averaging over all its possible realizations, we can quantify the amount of privacy loss and quality of service degradation. Further more, by assuming a simple point process model such as the Poisson point process, we can derive analytical results for the performance of several obfuscation mechanisms. These analytical results provide a first-order approximation for how an obfuscation mechanism will do and provide the designer with insight into the effect of varying the parameters of the mechanism.

3.6.1 Preliminaries

Before examining the analytical results, we must define a few key components of the stochastic geometry model. In this subsection, we define spatial point processes and explain how we adapt our privacy and quality of service metrics to the point process model.

Spatial Point Processes

We model these semantic locations as a finite set of points $\mathcal{S}$ in a bounded and connected region $\mathcal{X} \subset \mathbb{R}^2$. $\mathcal{S}$ could be a fixed point pattern, but more generally we model it a random, as a spatial point process, usually as a Poisson Point Process (PPP). A PPP on a set $\mathcal{X} \subset \mathbb{R}^2$ has to satisfy the following conditions:

- For any bounded $B \subseteq \mathcal{X}$, the number of points in B, denoted $N(B)$, is a Poisson random variable, and $E[N(B)] = \int_B \lambda(x)dx$, where $\lambda : \mathcal{X} \to \mathbb{R}_+$ is called the *density* of the process.
- If $B_1, ..., B_n$ are disjoint, bounded sets, then $N(B_1), ..., N(B_n)$ are independent random variables.

If the density function λ is constant over $\mathcal{X}$, then we say the PPP is *homogeneous*. If not, we say it is *inhomogeneous*.

You can also go a step further and define λ as a random process itself. In this case, the same conditions apply except the expectation of $N(B)$ becomes $E[N(B)] = \int_B E[\lambda(x)]dx$ [51]. Cox processes are used in practice to model a wide variety of point patterns due to their flexibility and computational tractability [35].

Alternate Characterization

When $\mathcal{X}$ is bounded, there is a simpler/more concrete alternative way of defining PPPs. Then $\mathcal{S}$ is a finite number N of points in $\mathcal{X}$, where N is a Poisson random variable with mean $\int_{\mathcal{X}} \lambda(x)dx$. Given N, $\{X_1, ..., X_N\}$ are i.i.d. with density

$$f_X(x) = \frac{\lambda(x)}{\int_{\mathcal{X}} \lambda(x)dx}.$$

In the homogeneous case, $\lambda(x) = \lambda$ and $f_X(x) = \lambda/(\lambda|\mathcal{X}|) = 1/|\mathcal{X}|$. A specific realization of a PPP can be denoted $\mathcal{S} = (n, \{x_1, ..., x_n\})$.

The privacy metric

We use Expected Success Rate (ESR) as the metric for location privacy. Informally, ESR is the probability that a strategic adversary (more on adversaries in the next section) can correctly infer your location. Equivalently, it is the expected proportion of semantic locations that can be correctly inferred. We would expect this to be 1 if users do not obfuscate their locations and expect it to be very low (but nonzero) for very heavy obfuscation, e.g. always reporting the same location.

It can also be thought of as an indirect way of measuring the value of the location data set to the adversary. As previously stated, semantic correctness is the expected proportion of semantic locations correctly inferred. But, importantly, the adversary doesn't know which ones are the correct ones. That means that if the ESR is, say, 95%, then the incorrect inferences may be an acceptable level of noise and the data set can be used (or abused) as intended. If, however, the ESR is 25%, then the dataset is of much less use. The adversary might be

inclined to discard the data (or not try to collect it in the first place). Information brokers might not want to buy the data knowing that it is more noise than signal. All in all, it could make the data less likely to be collected, stored, and used.

How do we measure ESR when $\mathcal{S}$ and π are stochastic? It is the same as before, $P(X = \hat{X})$, but this time the probability is not just over the randomness in the privacy mechanism K, but also in the realizations of $\mathcal{S}$ and π. To begin with, we make the simplifying assumption that π is uniform over $\mathcal{S}$. This allows us to derive analytical expressions for ESR for simple location privacy mechanisms such as uniform quantization and adding Gaussian or Laplacian noise. However, a uniform prior is likely to be inaccurate. We saw in section 3.5.3 that priors estimated from social media data is better modeled by a power-law distribution.

The Quality of Service metric

As for quality of service degradation (QoSD) metrics, both (α, δ)-usefulness and mean-squared error extend naturally from physical location privacy to semantic location privacy. We will, however, modify them slightly to give them a more intuitive interpretation. We'd like the QoSD to be proportional to the area of uncertainty, the size of the area where the user could possibly be. For many location-based services, the amount of extra bandwidth, memory, and processor resource consumption can reasonably be assumed to be proportional to the size of that area. For example, if the LBS is a turn-by-turn navigation app and the user sends a location report that indicates that the user could be anywhere inside an area A, the only recourse is to send the map tiles, routing info, and traf-

fic updates for the entire area. The bigger the area, the higher the bandwidth, memory, and CPU consumption.

For deterministic mechanisms such as quantization, the QoSD would simply be the size of the quantization bin. For probabilistic mechanisms such as adding noise, we can modify (α, δ)-usefulness slightly. We measure the size of the smallest area A such that $K(x)$ is within A with probability at least δ. For concreteness, we will fix $\delta = 0.9$ for the remainder of the chapter. Let the size of A be denoted by β. β can be derived from α because they are related by the formula for the area of a disc: $\beta = \pi\alpha^2$.

Adversarial Model

Let us make explicit the assumptions we have made about the adversary. The adversary knows $\mathcal{X}$, $\mathcal{S}$, π, and K. The adversary's goal is to correctly infer the semantic location of the user given their reported location. The adversary is *strategic* and therefore employs Bayes-optimal inference under a zero-one loss function (i.e. MAP inference).

3.6.2 Analytical Results

We are now ready to dive into the analysis. We have closed-form expressions that clearly show the effect of changing privacy mechanism parameters on location privacy as well as the effect of semantic location density (i.e. λ). We derive these results under rather strong assumptions, but as the next section will show, the insights derived from these results still hold given more sophisticated mod-

els (or with relaxed assumptions). Throughout this section, we will assume that the set of semantic locations, $\mathcal{S}$, is a homogeneous Poisson point process with intensity λ and that the prior, π, over them is the uniform distribution. We will show in the next section how to relax these assumptions at the cost of losing analytical tractability.

Quantization

Let us first consider the case where K is a uniform quantization mechanism. It turns out we can figure out a closed-form expression for ESR as a function of the quantization parameter Δ and semantic location density λ. For simplicity of calculations, we assume that $\mathcal{X}$ is a square region in $\mathbb{R}^2$ whose side lengths are an integer multiple of Δ.

It is possible to set up a simulation to compute ESR as a function of density λ and quantization parameter Δ. There is to the best of my knowledge no known derivation of this result, but it turns out that simulation agrees perfectly with the following expression:

$$\text{ESR}(\lambda, \Delta) = \frac{1 - e^{-\lambda\Delta^2}}{\lambda\Delta^2} \tag{3.4}$$

Because the derivation is not known, this result remains a conjecture at this stage. As for QoSD, we simply use the area of each bin as the QoSD metric. In other words,

$$\text{QoSD}(\Delta) = \Delta^2$$

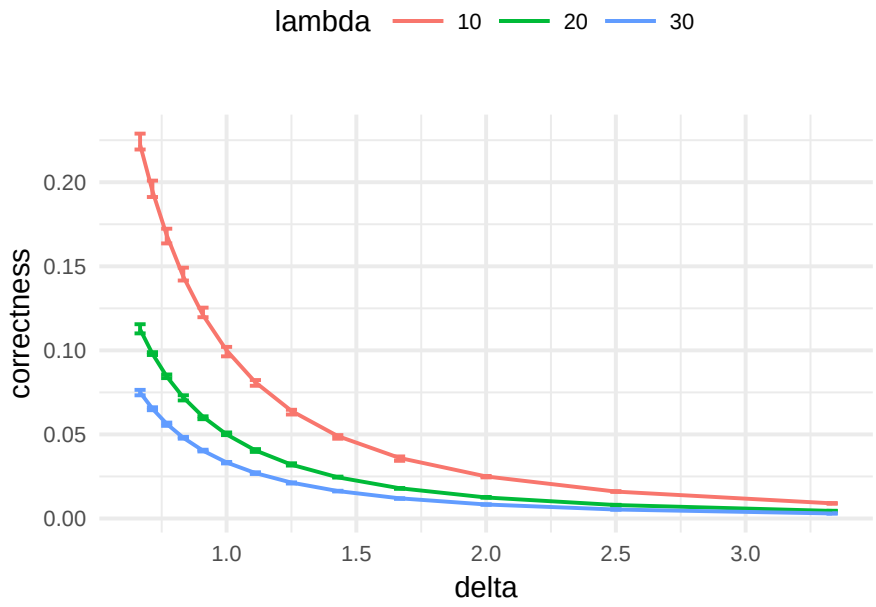


Figure 3.8: Simulation results verifying equation 3.4. The solid lines are the predicted values, while the points with error bars are the simulated results.

Additive Gaussian Noise

We can also derive a closed-form result for the ESR and QoSD when the privacy mechanism adds Gaussian noise with variance σ^2. We assume that $\mathcal{S}$ on $\mathcal{X} = \mathbb{R}^2$ is a homogeneous PPP and X an arbitrary point in $\mathbf{X}$. The reported location is $Y = X + W$, where $W \sim \mathcal{N}(0, \sigma^2 I)$ is zero-mean bivariate Gaussian with uncorrelated components. Because this point process is over all of $\mathbb{R}^2$, it has an infinite number of points with probability one. It is therefore ill-defined to say that π is uniform over $\mathcal{S}$. Instead, we will have to resort to an improper prior where $\pi(x) = 1$ for all $x \in \mathcal{S}$. Now, given a location report y, the Bayes-optimal decision under a uniform prior and additive, uncorrelated Gaussian noise is the nearest neighbor:

$$
\begin{aligned}
\hat{X}(y) &= \arg\max_{x \in \mathcal{S}} f_{X|Y}(x|y) \\
&= \arg\max_{x \in \mathcal{S}} \frac{f_{Y|X}(y|x)\pi(x)}{f_Y(y)} \\
&= \arg\max_{x \in \mathcal{S}} f_{Y|X}(y|x) \\
&= \arg\max_{x \in \mathcal{S}} f_W(y-x) \\
&= \arg\max_{x \in \mathcal{S}} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{||y-x||^2}{2\sigma^2}\right) \\
&= \arg\min_{x \in \mathcal{S}} ||y-x||
\end{aligned}
$$

To derive the ESR, we restate the problem slightly. Note that $B(Y, ||W||)$, the open ball of radius $||W||$ centered on Y, contains no points, since $\partial B(Y, ||W||) \backslash \{X\}$, the boundary of the ball minus X, almost surely contains no points. Therefore, $P(\hat{X} = X) = P(N(B(Y, ||W||) = 0)$. The radius $||W||$ follows a Rayleigh distribution with parameter σ (by the definition of Rayleigh RVs). For simplicity of notation, let's define $R := ||W||$. Then we are ready to compute the ESR:

$$
\begin{aligned}
P(\hat{X} = X) &= P(N(B(Y, R)) = 0) \\
&= \int_0^\infty P(N(B(Y, R)) = 0 | R = r) f_R(r) dr \\
&= \int_0^\infty \exp(-\lambda\pi r^2) \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) dr \\
&= \frac{1}{\sigma^2} \int_0^\infty \exp\left(-\left[\lambda\pi + \frac{1}{2\sigma^2}\right] r^2\right) r \, dr \\
&= \frac{1}{2\pi\lambda\sigma^2 + 1}
\end{aligned}
$$

And so we have our second result:

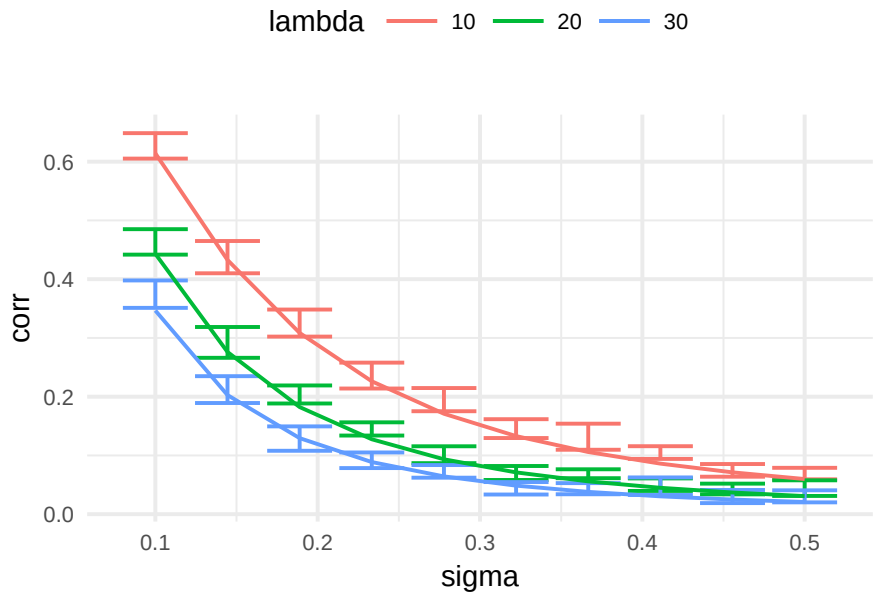


Figure 3.9: Simulation results verifying equation 3.5. The solid lines are the predicted values, while the points with error bars are the simulated results.

$$\text{ESR}(\sigma, \lambda) = \frac{1}{2\pi\lambda\sigma^2 + 1} \tag{3.5}$$

To derive the QoSD, we need to figure out for which α we have $P(||W|| \leq \alpha) = 0.9$. The Rayleigh quantile function tells us that $\alpha = \sigma\sqrt{2\log(10)}$. Thus the QoSD is

$$\begin{aligned} \text{QOSD}(\sigma) &= \pi\alpha^2 \\ &= 2\pi\sigma^2\log(10) \end{aligned}$$

Laplacian Noise

Using the same approach, we can derive a result for additive Laplacian noise. Laplacian noise is often of particular interest in privacy applications due to its connection to differential privacy. We use the planar Laplacian noise described by equation 3.1. The derivation is the same as for Gaussian noise, except W follows a planar Laplacian distribution with parameter σ and therefore $R = ||Z||$ has the distribution $f_R(r) = \frac{r}{\sigma^2}\exp(-r/\sigma)$. We derive the ESR:

$$\begin{aligned}
P(\hat{X} = X) &= P(N(B(Y,R)) = 0) \\
&= \int_0^\infty P(N(B(Y,R)) = 0 | R = r) f_R(r) dr \\
&= \int_0^\infty \exp(-\lambda\pi r^2)\frac{r}{\sigma^2}\exp(-\frac{r}{\sigma})dr \\
&= \frac{1}{\sigma^2}\int_0^\infty r\exp(-\lambda\pi r^2 - r/\sigma)dr \\
&= \frac{1}{2\pi\lambda\sigma^2} - \frac{1}{4\pi\sigma^3\lambda^{3/2}}\exp\left(\frac{1}{4\pi\lambda\sigma^2}\right)\text{erfc}\left(\frac{1}{2\sqrt{\pi\lambda}\sigma}\right)
\end{aligned}$$

So the ESR is therefore

$$\text{ESR}(\sigma,\lambda) = \frac{1}{2\pi\lambda\sigma^2} - \frac{1}{4\pi\sigma^3\lambda^{3/2}}\exp\left(\frac{1}{4\pi\lambda\sigma^2}\right)\text{erfc}\left(\frac{1}{2\sqrt{\pi\lambda}\sigma}\right) \tag{3.6}$$

To compute the QoSD, we once again need the quantile function for the distribution of the noise magnitude. $||W||$ follows the Gamma distribution with shape parameter fixed to $k = 2$ and scale parameter σ. As such, the QoSD does not have a closed-form expression, we can be evaluated numerically as follows:

$$\text{QoSD}(\sigma) = \pi\alpha^2$$

where

$$\begin{aligned}\alpha &= Q(0.9; 2, \sigma) \\ &= \sigma\gamma^{-1}(2, 0.9)\end{aligned}$$

where $Q(p; k, \sigma)$ is the quantile function of the Gamma distribution with shape parameter k and scale parameter σ and $\gamma^{-1}(k, y)$ is the inverse of the lower incomplete gamma function defined as follows:

$$\gamma(k, x) = \int_0^x t^{k-1} e^{-t} dt$$

Using numerical evaluation, we can determine that

$$Q(0.9; 2, \sigma) \approx 3.89\sigma$$

and therefore

$$\text{QoSD}(\sigma) \approx 15.1\pi\sigma^2$$

3.6.3 Conclusion

Deriving closed-form expressions of ESR based on stochastic geometry models is a difficult problem. I was able to derive results using very simple models. Those models relied on two rather unrealistic assumptions. They assumed the

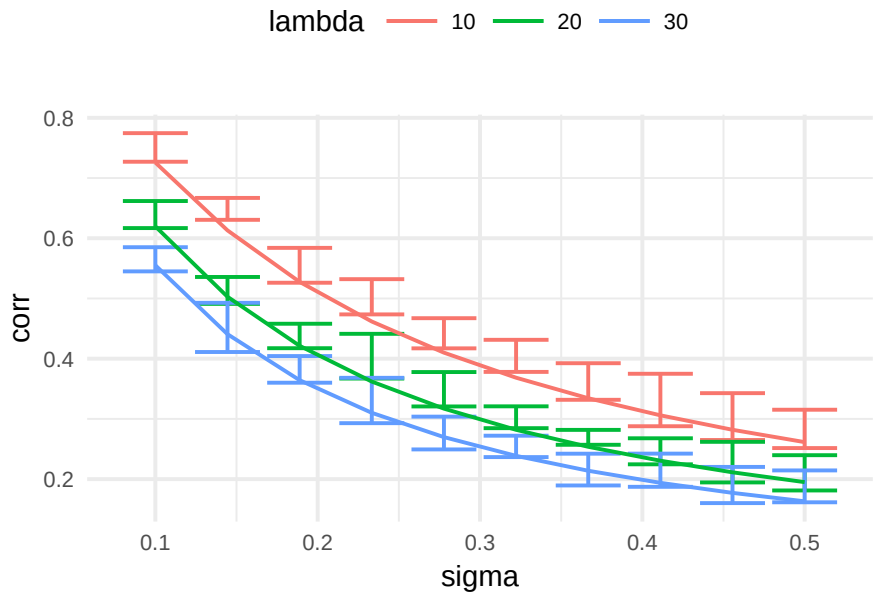


Figure 3.10: Simulation results verifying equation 3.6. The solid lines are the predicted values, while the points with error bars are the simulated results.

spatial distribution of semantic locations was that of a point process. Based on the location-based social media data gathered, that assumption appears highly unlikely. The point patterns of semantic locations exhibited a high variability of density. Future research could examine the viability of modeling said point patterns with more sophisticated point processes, such as the log-Gaussian Cox process, the Thomas process, Matern Cluster process, or a Neyman-Scott process.

The other assumption I relied on was that the prior over the semantic locations is uniform. This could be a more problematic assumption. The research in section 3.5 clearly showed that popularity of semantic locations in location-based social media follows a power law distribution. Deriving closed-form expressions under a power law distributed prior, e.g. Zipfian prior, could greatly increase the impact and utility of this line of research.

3.7 Differential Privacy Analysis

In this section, we extend differential privacy to semantic location privacy. This allows us to make semantic location privacy guarantees that are independent of the adversary's prior. As explained in previous sections, differential privacy has already been adapted to physical location privacy under the name geo-indistinguishability. Here, we will show that geo-indistinguishability is not an appropriate metric for semantic location privacy. To remedy that, we define a new version of geo-indistinguishability, called *semantic geo-indistinguishability* (SGI).

3.7.1 Shortcomings of Geo-indistinguishability

When trying to protect physical location privacy, geo-indistinguishability is a perfectly reasonable approach. If, however, we care about semantic location privacy, it may not measure privacy the way we expect. If we're in an area where semantic locations are few and far between, ϵ-geo-indistinguishability may offer a false sense of security. Let's consider an example where we make the optimistic assumption that all semantic locations are equally likely a prior. Once again, let $\mathcal{S}$ be the set of semantic locations under consideration, and $x \in \mathcal{S} \subset \mathcal{X}$ be the true location of the user. Then their report is simply the true location with added noise, $Y = K(x) = x + W$, where W is the planar Laplace distribution. The planar laplace distribution with scale parameter σ is given in polar coordinates as (R, Θ), where $R \sim \text{Gamma}(2, \sigma)$ and $\Theta \sim \text{Unif}(0, 2\pi)$. We want to provide ϵ'-privacy within a radius r, so $\epsilon = \epsilon'/r$, and thus the noise added has scale parameter $\frac{1}{\epsilon} = \frac{r}{\epsilon'}$. The Bayes-optimal inference under zero-one

loss and Laplace-distributed noise is to infer

$$\hat{X} = \arg\min_{x \in \mathcal{S}} ||x - Y||$$

i.e. $\hat{x}$ is the closest point in $\mathcal{S}$ to y. But what if x is the closest semantic locations to all points in $B(x, r)$? Then the ESR is

$$\begin{aligned} P(\hat{X} = x) &\geq P(Y \in B(x, r)) \\ &= P(||W|| < r) \\ &= 1 - \left(\frac{r}{\sigma} + 1\right) \exp\left(-\frac{r}{\sigma}\right) \\ &= 1 - (\epsilon r + 1) \exp(-\epsilon r) \\ &= 1 - (\epsilon' + 1) \exp(-\epsilon') \end{aligned}$$

If we set ϵ' to the relatively low value of $\log 5$, then the ESR is $P(\hat{X} = x) \gtrapprox 0.5$. Thus the ESR is greater than 50%, even though we have provided geo-indistinguishability with a relatively low ϵ' and even if we assume no prior information on behalf of the adversary. The problem here is that geo-indistinguishability is only an indirect measure of semantic location privacy. In the next section, we propose a modification of differential location privacy that makes semantic locations a first-class citizen.

3.7.2 Semantic Geo-indistinguishability

We propose a definition of differential location privacy that takes semantic locations into account. Geo-indistinguishably provides differential privacy for all

physical locations within a set $B(x, r) \in \mathcal{X}$. Semantic geo-indistinguishability aims to provide differential privacy within a set $R \subseteq \mathcal{S}$. The definition is as follows. A privacy mechanism K provides ϵ-semantic differential privacy within R if for all $x, x' \in R$ and all $S \subseteq \text{range}(K)$,

$$P(K(x) \in S) \leq e^{\epsilon} P(K(x') \in S)$$

3.7.3 Properties of Semantic Geo-indistinguishability

Composition Theorem

It is easy to show that the celebrated composition theorem for differential privacy carries over to semantic geo-indistinguishability. Informally, the composition theorem says that releasing information from two mechanisms that provide SGI also provides SGI whose SGI level is the sum of the individual SGI levels of the two mechanisms. I.e. if K_1 provides ϵ_1 SGI and K_2 provides ϵ_2 SGI, then (K_1, K_2) provides $\epsilon_1 + \epsilon_2$ SGI.

Theorem 1. *Let $K_1 : \mathcal{X} \to \mathcal{Y}_1$ provide ϵ_1-SGI and $K_2 : \mathcal{X} \to \mathcal{Y}_2$ provide ϵ_2-SGI. Then define $K_{1,2} : \mathcal{X} \to \mathcal{Y}_1 \times \mathcal{Y}_2$ with the mapping $K_{1,2}(x) = (K_1(x), K_2(x))$. Then $K_{1,2}$ provides $(\epsilon_1 + \epsilon_2)$-SGI.*

Proof. Let (y_1, y_2) be an arbitrary point in $\mathcal{Y}_1 \times \mathcal{Y}_2$ and let x, x' be arbitrary points

in $\mathcal{X}$. Then

$$\begin{aligned} P(K_{1,2}(x) = (y_1, y_2)) &= P(K_1(x) = y_1 \wedge K_2(x) = y_2) \\ &= P(K_1(x) = y_1)P(K_2(x) = y_2) \\ &= e^{\epsilon_1} P(K_1(x') = y_1) e^{\epsilon_2} P(K_2(x') = y_2) \\ &= e^{\epsilon_1 + \epsilon_2} P(K_1(x') = y_1) P(K_2(x') = y_2) \\ &= e^{\epsilon_1 + \epsilon_2} P(K_{1,2}(x') = (y_1, y_2)) \end{aligned}$$

□

Max Divergence Interpretation

SGI also retains the interpretation of DP as Max Divergence. Max Divergence is defined as follows:

Definition 1. *Let X_1 and X_2 be random variables with the same support $\mathcal{X}$. Then Max Divergence is given by*

$$D_\infty(X_1 || X_2) = \max_{S \subseteq \mathcal{X}} \left| \log \frac{P(X_1 \in S)}{P(X_2 \in S)} \right|$$

It is easy to show that a mechanism K provides ϵ-SGI if and only if $D_\infty(K(x)||K(y)), D_\infty(K(y)||K(x)) \leq \epsilon$.

3.8 Randomized response mechanism

This section introduces a randomized response mechanism that provides semantic differential privacy while preserving quality of service for users and preserving the utility of aggregate statistics of location information. This mechanism is inspired by the RAPPOR mechanism [42], a variation of the classical randomized response mechanism introduced in [128].

The Mechanism

Let $R_1, ..., R_M$ form an equisized partition of a set of semantic locations $\mathcal{S}$. Let $|R_1| = ...|R_M| = N$. The randomized response algorithm provides privacy by reporting a set of locations instead of simply a single one. The set of reports can be represented as a binary vector $b = (b_1, ..., b_N)$, where $b_i = 1$ means that location $x_i \in R_k$ was included in the set. The values of b are set according to the following randomized method. Let x_j be the true location. Then $b_i \sim$ Bernoulli(p) if $i = j$ and $b_i \sim$ Bernoulli(q) if $i \neq j$. The report sent to the LBS will be (R_k, b).

3.8.1 Analysis

Let b be a candidate response. Without loss of generality, we can rearrange b such that $b = (b_1, ..., b_h, b_{h+1}, ..., b_N)$, where $b_1, ..., b_h$ are all ones and $b_{h+1}, ..., b_N$ are all zeros. Then for any $x \in R$,

$$P(K(x) = b) = \begin{cases} pq^{h-1}(1-q)^{N-h} & x \in \{x_1, ..., x_h\} \\ q^h(1-p)(1-q)^{N-h-1} & x \in \{x_{h+1}, ..., x_N\} \end{cases}$$

The differential privacy parameter is given by the following equation:

$$\epsilon = \max_{x,x',b} \log \frac{P(K(x) = b)}{P(K(x') = b)}$$

Assuming $p \geq q$, the ratio is maximized when $x \in \{x_1, ..., x_h\}$ and $x' \in \{x_{h+1}, ..., x_N\}$ and is then

$$\begin{aligned} \frac{P(K(x) = b)}{P(K(x') = b)} &= \frac{pq^{h-1}(1-q)^{N-h}}{q^h(1-p)(1-q)^{N-h-1}} \\ &= \frac{p(1-q)}{q(1-p)} \end{aligned}$$

Thus the level of semantic differential privacy provided by this mechanism is

$$\epsilon = \log \frac{p(1-q)}{q(1-p)}$$

As a concrete example, let $p = 0.75$ and $q = 0.375$. Then $\epsilon = \log 5$.

Expected Success Ratio analysis

Let us consider the ESR using our scheme and compare it with geo-indistinguishability. Let $p = 0.75$, $q = 0.375$ for a differential privacy of $\epsilon = \log 5$.

Also, let $N = 100$, where N is the size of the partitions of the set of semantic locations. To simplify the notation in what follows, define D as the event that X is included in the location report. Like in the case of geo-indistinguishability, we assume that all semantic locations are equally likely a priori. Then we compute the ESR as follows:

$$P(\hat{X} = X) = P(\hat{X} = X|D)P(D) + P(\hat{X} = X|D^c)P(D^c)$$

We know that $P(D) = p = 1 - P(D^c)$. Then only two terms remain. Let us compute them one by one. We partition the event space on the number of locations included in the report, $\sum_{i=1}^{N} b_i$:

$$P(\hat{X} = X|D) = \sum_{k=1}^{N} P(\hat{X} = X|D, \sum_{i=1}^{N} b_i = k)P(\sum_{i=1}^{N} b_i = k|D)$$

Let us consider each term separately. Under a uniform prior, the Bayes-optimal inference is to pick an arbitrary location out of the k reported locations if $k > 0$ and pick an arbitrary location out of the N total locations if $k = 0$. Thus

$$P(\hat{X} = X|D, \sum_{i=1}^{N} b_i = k) = \frac{1}{k}$$

We can also see that the conditional distribution of $\sum_{i=1}^{N} b_i|D$ is basically a binomial random variable:

$$P\left(\sum_{i=1}^{N} b_i = k|D\right) = \binom{N-1}{k-1} q^{k-1}(1-q)^{N-k}$$

The other term we need to consider is $P(\hat{X} = X|D^c)$, the probability of the event that the adversary correctly infers the user's location given that it is not included in the location report. If the location report is not empty, then under Bayes-optimal inference, the adversary will not be correct, since they will choose of the locations in the location report. If the location report is empty, any semantic location in R is Bayes-optimal and therefore the probability of correctly inferring the user's location is $1/N$. Therefore,

$$\begin{aligned} P(\hat{X} = X|D^c) &= P(\hat{X} = X|\sum_{i=1}^{N} b_i = 0, D^c)P(\sum_{i=1}^{N} b_i = 0, D^c) \\ &= \frac{1}{N}(1-q)^{N-1} \end{aligned}$$

Putting all of this together, we get

$$\begin{aligned} P(\hat{X} = X) &= P(\hat{X} = X|D)P(D) + P(\hat{X} = X|D^c)P(D^c) \\ &= p\sum_{k=1}^{N} \frac{1}{k}\binom{N-1}{k-1} q^{k-1}(1-q)^{N-k} + (1-p)\frac{1}{N}(1-q)^{N-1} \end{aligned}$$

Evaluating that expression numerically for $p = 0.75, q = 0.375, N = 100$, we get

$$P(\hat{X} = X) \approx 0.02$$

This is independent of the density of semantic locations, since the set of locations is partitioned by semantic locations and not by area. Compare this with geo-indistinguishability also with $\epsilon' = \log 5$, where the corresponding ESR was 0.5.

Recovering aggregate statistics

The advantage of reporting the randomized response (R, b) instead of just reporting the region R is that this allows the LBS to recover aggregate statistics on mobility. We can recover aggregate statistics as follows. For a given location i, let $B_i := \sum_{j=1}^{N} b_i^{(j)}$, i.e. B_i is the total number of location reports that included location i (out of N reports). Taking the expectation of B_i and assuming location reports are independent, we get

$$\begin{aligned}
E[B_i] &= E\left[\sum_{j=1}^{N} b_i^{(j)}\right] \\
&= \sum_{j=1}^{N} E[b_i^{(j)}] \\
&= NE[b_i^{(1)}] \\
&= N\{E[b_i^{(1)}|t_i^{(1)} = 1]P(t_i^{(1)} = 1) + E[b_i^{(1)}|t_i^{(1)} = 0]P(t_i^{(1)=0})\} \\
&= N(p\pi_i + q(1 - \pi_i)) \\
&= \pi_i N(p - q) + Nq
\end{aligned}$$

Where $t_i^{(j)}$ is 1 or 0 depending on whether the user was truly at location i when making report j, and π_i is the probability that given a random location report, its user will be at location i. In other words, it is the true proportion of

location reports where the user is actually at location i. Rearranging the above equation, we get

$$\pi_i = \frac{E[B_i] + Nq}{N(p-q)}$$

Thus we can use the following estimator for π_i:

$$\hat{\pi_i} = \frac{B_i + Nq}{N(p-q)}$$

It can be shown that this estimator is unbiased, consistent, and asymptotically normal.

Bayesian Analysis

A Bayes-optimal adversary using the 0-1 loss function will employ the MAP rule for inference. Let us derive an expression for the MAP inference. Let the reported set be $y = \{x_1, ..., x_k\}$. Then

$$\begin{aligned}
\arg\max_{x \in \mathcal{X}} p(x|x_1, ..., x_k) &= \arg\max_{x \in \mathcal{X}} \frac{p(y_1, ..., y_k|x)\pi(x)}{\sum_{x' \in \mathcal{X}} p(y_1, ..., y_k|x')\pi(x')} \\
&= \arg\max_{x \in \mathcal{X}} p(y_1, ..., y_k|x)\pi(x) \\
&= \arg\max_{x \in \{x_1, ..., x_k\}} \pi(x)
\end{aligned}$$

Given this inference algorithm, what is the adversary's ESR? Let X be the user's true location and Y be the set of reported locations. Then

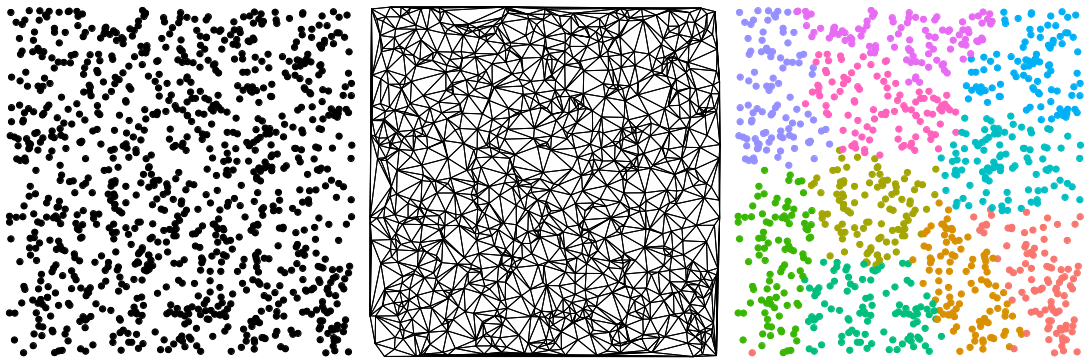
$$
\begin{aligned}
\text{ESR} &= P(X = \hat{X}) \\
&= P(X \in Y)P(X = \hat{X}|X \in Y) \\
&= p \cdot P(X = \hat{X}|X \in Y) \\
&= p\frac{\max_{x \in Y} \pi(x)}{\sum_{x \in Y} \pi(x)}
\end{aligned}
$$

The QoSD will be measured as the convex hull of the reported locations Y:

$$
\text{QoSD} = \text{Conv}(Y)
$$

3.8.2 Implementation Details

Ideally, we would find a set of equally sized partitions of C that would minimize both ESR and QoSD for the given partition size. This is a hard combinatorial problem, so we have devised a heuristic approximation. We construct a graph by computing the Delaunay triangulation for the set of semantic locations. We subsequently compute a balanced graph partition using the multilevel k-way partitioning scheme implemented in the METIS software package [68]. The METIS algorithm seeks to minimize the number of cut edges between roughly equisized partitions and will therefore tend to produce partitions that are geographically contiguous.



(a) (b) (c)

Figure 3.11: Illustration of the heuristic partioning algorithm. We construct a Delaunay graph (b) from the original point pattern (a). Then a balanced graph partioning algorithm is applied to the Delaunay graph, resulting in spatially homogeneous partitions (c).

3.8.3 Experimental Evaluation

To further evaluate our semantic geo-indistinguishability scheme, we perform a series of experiments using location-based social media data to construct sets of semantic locations and priors over them. Dataset citation: [139, 140]. This experiment is made using semantic locations from Boston, Massachusetts. For both regular geo-indistinguishability and semantic geo-indistinguishability, we use a differential privacy parameter of $\epsilon = \log 3$. We vary the parameters R, the GI radius, and N, the SGI partition size, and compute the quality of service degradation (QoSD) and ESR for that point. The result is a tradeoff curve between QoSD and ESR. This allows us to compare disparate privacy schemes.

Even though our randomized response mechanism is not uniformly better than the mechanism for regular geo-indistinguishability, it accesses a part of the trade-off space that the other cannot. The curve for SGI is not as smooth as the curve for GI for two reasons: SGI uses a heuristic algorithm to determine partitions as optimal graph partitioning is a hard combinatorial problem, and

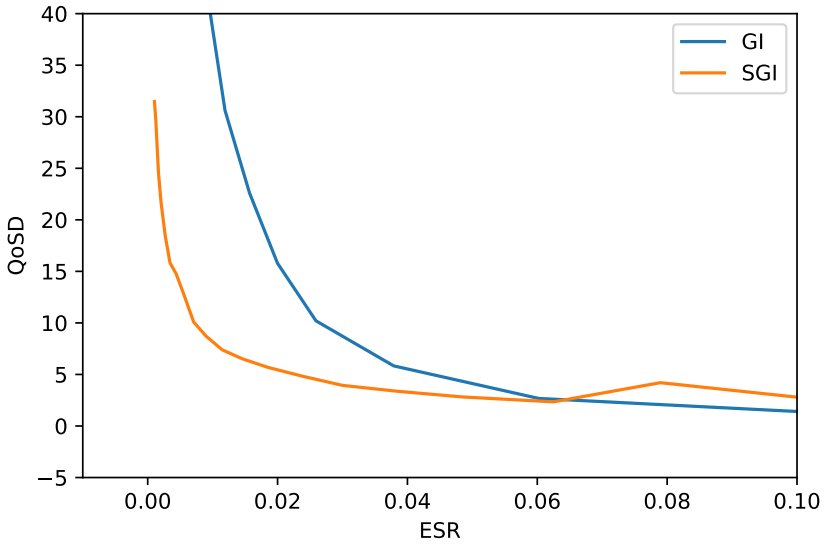


Figure 3.12: Temporary figure. Need to run simulation longer to get more accurate estimates and need to include error bars on simulated values. Shows rough tradeoff curves using SI and SGI using data from Boston. SGI clearly has a much better ”sweet spot”.

secondly, because the ESR and QoSD of SGI depends on partitioning a discrete, finite number of points, they cannot possibly be truly continuous functions.

3.9 Summary

In this chapter, I have explored the problem of location privacy in the Internet of Things through the lens of semantic location privacy. I reviewed the main existing approaches to building and analyzing location privacy mechanism. I went on to develop two separate approaches to analyzing semantic location privacy, an empirical approach and a theoretical framework. Lastly, I introduced a privacy mechanism of my own design and showed that a combination of Bayesian and Differential Privacy analysis provides a better picture than either one in isolation. Lastly, we presented an approach to analyzing and comparing mechanisms using privacy-utility trade-off curves. These curves allow us to directly compare two very different mechanisms and analyze them for a whole range of parameter choices simultaneously.

While it is true that semantic location privacy presents additional challenges compared to physical location privacy, semantic location privacy is still viable. Using social media data and stochastic models, we can concretely evaluate the engineering trade-offs of location privacy mechanisms. Furthermore, most of our analysis showed that by adding just a little bit of obfuscation, we achieve large privacy gains. This gives hope that a good trade-off between privacy and utility is possible.

The last take-away is that differential privacy analysis is not sufficient. In its strictest form, it is too conservative to yield anything useful. It is relaxed form, it is not sufficiently strong to provide meaningful privacy guarantees. A combination of differential privacy and Bayesian analysis is needed. Ideally, a location privacy mechanism will satisfy a relaxed form of differential privacy,

providing unconditional (but probabilistic) guarantees. The Bayesian approach can show us that given reasonable but strong assumptions on the adversary, we can still provide guarantees.

CHAPTER 4

VEGVISIR: AN IOT BLOCKCHAIN ENABLING PRIVACY-AWARE ACCESS TO MEDICAL RECORDS

They constantly try to escape
From the darkness outside and within
By dreaming of systems so perfect that no one will need to be good

T.S. Eliot

4.1 Introduction

In the previous two chapters we have considered two significant threats to privacy that have arisen in the IoT. In this chapter, we will see that the IoT doesn't have to be a threat to privacy, but can also be a boon to it. We present Vegvisir, a partition-tolerant blockchain that enables fast, reliable, and secure access to medical records for emergency first responders during disaster response.

Blockchains have emerged as an exciting new paradigm for distributed systems. From Bitcoin's cryptocurrency [93] to Ethereum's Turing-complete smart contracts [135], blockchains are being explored as a means to solve problems across a wide range of industries, including banking, energy, transportation and accounting. Blockchains could have important uses in IoT and Edge Cloud environments as well [26]. Such systems are often used for applications that require tamperproof logging of events for accountability. They are deployed under varied administrative domains, and so the decentralized ownership of blockchains make them a good match. However, current blockchain designs require high

network connectivity and are power-intensive, both of which can detract from their utility in such environments.

A blockchain is simply a tamperproof log of transactions. Blockchain implementations use a distributed trust model, removing the need for centralized control and single-point-of-failure designs. As long as a large enough fraction of participants execute the protocol (usually half or one-third), its security properties will be enforced. This provides a system that is both strongly consistent and highly available. It is not, however, tolerant of network partitions. Partitions cause branches (AKA forks) in the blockchain and branches must be resolved, meaning only one branch gets to stay a part of the blockchain while all others are discarded. With network partitions, such branches may stay around for a long time and lead to undesirable behaviors, even if branches are eventually resolved. For example, people who have conducted business with Bitcoin may find that the bitcoins they were paid are now back in the hands of the original owner or have been spent otherwise. Additionally, most current blockchain designs are very energy-intensive, requiring vast amounts of computation solving cryptopuzzles. Bitcoin alone is estimated to use tens of terawatt hours per year, enough to power a mid-sized country [98, 9].

These two characteristics, the need for high network reliability and high power consumption, make Bitcoin and most other existing blockchain designs unsuitable for deployment in ad hoc IoT networks or edge cloud systems. We present *Vegvisir*, a blockchain specifically designed for the low-connectivity, low-power IoT setting. It tolerates network partitions well and uses a low-power consensus mechanism. Instead of resolving branches, it permits them, resulting in a Directed Acyclic Graph (DAG) structure of the blockchain rather

than a linear one. The cost of this partition tolerance is that the types of applications that can be implemented with the blockchain are limited to ones that only require a partial ordering of logged events. To this end, Vegvisir supports applications based on Conflict-free Replicated Data Types (CRDT) [113].

To motivate the need for Vegvisir, we present a use case for it in the area of disaster response in Section 4.2. In section 4.3, we review some foundational concepts from cryptography and distributed systems on which Vegvisir is based. We go on to survey related literature in section 4.4 and develop the architecture of the blockchain in section 4.5. Section 4.6 describes our initial implementation of the blockchain and section 4.7 discusses the implications and challenges of this line of research.

The work in this chapter will appear in the proceedings of the 38th IEEE International Conference on Distributed Computing Systems under the name *Vegvisir: A Partition-Tolerant Blockchain for the Internet-of-Things*.

4.2 Motivation

The 2017 Atlantic hurricane season was one of the worst on record. Three major hurricanes devastated the Caribbean, Florida, and Texas. Hundreds of people lost their lives and the property damage is estimated to be over $300 billion [104]. The loss of lives, limbs, and property had undoubtedly been greater if not for the valiant efforts of thousands of emergency first responders. If first responders could leverage more information and communication technology to aid and coordinate their efforts, further lives could potentially be saved. In an ideal world, first responders have a strong communication network and a ro-

bust cloud infrastructure that enables information to flow to the right places at the right time and eases coordination of rescue efforts. Natural disasters, however, can render communication infrastructure such as cell towers and Land Mobile Radio System (LMRS) repeaters inoperable. First responders must in those cases deploy their own communication infrastructure as well as take advantage of every possible means of communicating, forming heterogeneous ad hoc mobile networks to make up for lack of connectivity. Existing communication and cloud infrastructure is not built to operate in such environments. We need a new infrastructure design to enable first responder applications.

While cellular phones may also end up inoperable after extended periods of time in a disaster area, it is important to maximize the use of resources as long as they are available. To that end, device-to-device (D2D) communication for public safety scenarios in LTE and 5G networks has received considerable attention (see for example [43, 126, 36, 138]). Our work builds on such efforts and provides distributed applications implemented on top of unreliable, ad hoc networks such as the D2D networks mentioned above.

One of the problems medical personnel face both in and outside of emergency situations is the need for accessing electronic health records promptly while safeguarding their security and privacy. We propose that blockchains can be used to implement a use-based privacy solution that gives emergency first responders ready access to sensitive patient health records but enforces strict accountability. Use-based privacy is an approach to privacy that focuses on uses (and abuses) of sensitive records, rather than access [20] and has in recent years been proposed as a framework under which to design privacy policies [21, 14]. Patients generally will not object to a physician or paramedic accessing their

medical records in order to help save their lives (a valid use) but they would object to the same physician accessing their records without a medical reason. During emergencies, paramedics and physicians could have all their access requests to sensitive records granted under the condition that the request has been recorded in a tamperproof log. Once the state of emergency is over, the log is reviewed. If frivolous access has occurred, such as a medical worker accessing an ex-spouse's or a celebrity's health record, the worker could be sanctioned, providing incentive to only access health records when necessary.

Our approach with Vegvisir presents a good avenue to implement a tamperproof log in such an environment. It consists of an unreliable network between many low-power IoT devices (first responder smartphones), some of whom cannot be fully trusted. Our solution can ensure that no health record is accessed without an explicit request for access being persistently stored on the blockchain. It does not require proof-of-work and is therefore easy on the batteries, and its opportunistic gossip-style protocol for spreading blocks is well-suited for a mobile ad hoc network.

4.3 Preliminaries

In this section, we review the basics of the key technologies underlying Vegvisir. Vegvisir is at its core a blockchain, a distributed cryptographic data structure that provides extremely high fault tolerance. Blockchains, in turn, are based on the cryptographic primitives of digital signatures and cryptographic hash functions. Unlike other blockchains presented to date, Vegvisir is also an implementation of a type of data structure called a Conflict-free Replicated Data Type

(CRDT). We will review CRDTs and their implications for the cornerstone theorem of the field of distributed systems, the CAP theorem. Finally, as Vegvisir takes an approach to security and privacy that relies on accountability more than access control, we will review the basic tenets of security engineering as they relate to those concepts.

4.3.1 Cryptographic Primitives

There are two main cryptographic primitives used in the Vegvisir blockchain: *cryptographic hash functions* and *digital signatures*. These primitives are crucial to ensuring tamperproofness, data provenance, and authentication in the blockchain. The definitions in this subsection are adapted from [87].

Cryptographic Hash Functions

Cryptographic hash functions are the workhorse of modern cryptography. A cryptographic hash function h takes as input a binary string of arbitrary length and output a binary string of a fixed length:

$$h : \{0,1\}^* \rightarrow \{0,1\}^n$$

where $\{0,1\}* = \bigcup_{k=0}^{\infty}\{0,1\}^k$. Both cryptographic and non-cryptographic hash functions are generally designed to be easy to compute. Cryptographic hash functions need to additionally satisfy the properties of preimage resistance, second preimage resistance, and collision resistance. A hash function is said to

have preimage resistance if for any specified output $y \in \{0,1\}^n$, it is computationally infeasible to find an input $x \in \{0,1\}^*$ such that $h(x) = y$. It is said to have second preimage resistance if given an input x, it is computationally infeasible to find a second input x' such that $h(x) = h(x')$. Collision resistance implies that it is computationally infeasible to find a pair of inputs x, x' such that $h(x) = h(x')$. In addition to those properties, practical cryptographic hash functions are also desired to have the informal property of the *avalanche effect*. The avalanche effect implies that small changes in the input, such as changing a single bit, will result in drastic changes in the output.

We leave the terms "computationally infeasible" and "easy to compute" purposefully undefined, as practical hash functions, such as the Secure Hash Algorithm (SHA) family of hash functions [32, 40], have no theoretical proofs that they satisfy any reasonable definition.

Digital Signatures

Digital signatures use asymmetric cryptography to ensure authenticity of messages. A user will generate a *signing key* K_S and keep it secret, as well as a *verification key* K_V, which is made public. The signing function will take as input a string of arbitrary length (the message) and a signing key of fixed length and produce a tag. The verification function will take the message, the tag, and the corresponding verification key as input and produce 0 or 1 as output. If it produces 1 as output, then the tag was indeed the result of applying the signing function to the message with the corresponding signing key.

Blockchains

Blockchains are a cryptographic technology first introduced in 2008 in a white paper describing the first cryptocurrency, Bitcoin [93]. They underlie all modern cryptocurrencies and could potentially be useful in a broad range of other applications. Their tamperproof data structures with clear data provenance and distributed trust model could make them useful in banking, supply chain management, auditing, and potentially many more industries.

The key aspect of blockchains is that they enable actors in a distributed system to agree on the contents of a shared data structure, usually called a ledger due to blockchains' financial roots, without fully trusting one another. This allows participants in a blockchain network to agree on the order and content of a series of transactions, which form the basis for a currency.

Blockchains are a family of technologies that enable a distributed network of actors to agree on the contents and order of a shared data structure without relying on a central authority. The first blockchain was introduced by Satoshi Nakamoto in 2008 [93] as a means of implementing a decentralized digital currency. The blockchain contains a series of *blocks*, starting with the *genesis block*. Each block consists of a *block header* and a series of *transactions*. The transactions would be along the lines of "Transfer x bitcoin from account A to account B" and would have to be signed by the the signing key associated with account A. The block header contains some metainformation, the most important of which are the *previous block hash* and the *nonce*. The previous block hash is the hash of the previous block as computed by a specified hash function. The nonce is a number that, once added to the blockchain header, ensures that the first k bits of the blockchain hash are zero. The only known way to find such a nonce is by

brute force. This is called a *proof-of-work*, and is the cornerstone in the Bitcoin security model.

The proof-of-work is set to a difficulty level where the entire network generates a new block with a valid nonce every 10 minutes. Therefore, if a malicious actor in the blockchain network wants to change a block already in the blockchain, they would have to find a new nonce not only for the block they modified, but all subsequent blocks. This is because the blocks are chained together by including the hash of the previous block in each block.

Permissioned vs public blockchains

Public blockchains such as Bitcoin and Ethereum allow anyone to participate in the blockchain protocol provided they appear to be following the protocol. Their security rests on the computational difficulty of a cryptographic puzzle called *proof-of-work*. A typical puzzle will require the user to partially invert a cryptographic hash function. This is a hard problem by design. It requires a brute-force approach, i.e. trying different inputs until the desired output is achieved. This approach is computationally intensive and therefore consumes a lot of power.

Permissioned blockchains, on the other hand, can avoid these computationally expensive operations. In a permissioned blockchain, a central authority decides who can and cannot participate in the network. This reduces the decentralization of the blockchain, but makes it much less computationally demanding.

4.3.2 CAP Theorem

The CAP theorem is one of the most important theoretical results in distributed computed. It concerns the tradeoffs between *consistency*, *availability*, and *partition tolerance* in a distributed system. A distributed system consisting of multiple nodes is said to be consistent if all nodes agree on the state of the system. It is said to be available if it responds to requests/agitation? with a reasonable time frame (this is not a good definition). And it is said to be partition-tolerant if it continues to function even if there is not a *network path* between all pairs of nodes. The celebrated CAP theorem states that a distributed system can at most satisfy two out of those three properties [53].

4.3.3 Conflict-Free Replicated Data Types

Conflict-free Replicated Data Types are data structures for distributed systems that can be updated independently and concurrently by multiple replicas in a distributed system and subsequently merged without any conflicts in a principled manner. There are two equivalent formulations of CRDTs. The first one, often called operation-based, requires the data structure to be equipped with a commutative update function. The second one, state-based CRDTs, require a merge operation that is commutative, associative, and idempotent.

The advantage of CRDTs in distributed systems is that they enable multiple replicas in a system to operate independently and concurrently on the same data structure, but still provide unambigious semantics when it comes to merging. Recasting blockchains as a CRDT is one of the key contributions of this chapter.

4.3.4 The Golden Rule

In the field of systems security, the Golden Rule refers to the three essential components to achieving confidentiality and privacy of information: (au)thentication, (au)thorization, and (au)dit. Authentication means that the system verifies that the user asking for access is indeed the user they claim to be. Authorization means that the system verifies that said user is permitted to access the data in question. Audit means that a user's access to said record is recorded in a log that can later be inspected. Audits allow users to be held accountable for their actions if any wrongdoing is suspected.

Computer security and privacy mechanisms tend to focus a lot on the authentication and authorization components, but a lot less on the audit components. Audit and accountability can be a powerful deterrent and is underused in many systems. Vegvisir will rely heavily on accountability to enforce its privacy properties.

4.4 Related Works

Blockchains were first introduced in 2008 as part of the then novel Bitcoin cryptocurrency system [93]. Since then, the blockchain field has seen explosive growth with many variants and use cases proposed. One of the more notable variants is Ethereum, which replaces the basic scripting language implemented in Bitcoin with a Turing-complete one, paving the way for so-called smart contracts [135]. Both of these blockchains have a linear structure and rely on a proof-of-work consensus mechanism which requires solving a computationally

expensive cryptopuzzle, making them poorly suited to our use cases.

Many variants on the Bitcoin protocol have been proposed since then, some of which use a DAG structure like Vegvisir. The GHOST protocol is a modification to the Bitcoin blockchain that uses a DAG structure to improve security [120]. The point of this modification is to enable a more robust method of selecting which fork to keep and which to discard. By keeping track of all forks, a node can choose a fork based on the heaviest-subtree-wins rule (the subtree with the largest number of blocks) as opposed to the longest-chain-wins rule, wasting less work and thus eliminating certain forms of attacks.

The Byteball blockchain platform proposes a new type of cryptocurrency with a DAG structure [28]. Byteball eliminates the distinction between blocks and transactions. Each 'block' is a single transaction and can have multiple parents. Double spending is prevented by determining a total order on the DAG through the behavior of a set of privileged users called 'witnesses'. The total order is used to determine which transactions to keep and which to declare invalid when double spending occurs.

Iota is perhaps the best-known implementation of a blockchain with a DAG structure [106]. Iota is a transaction fee-less cryptocurrency where double spends are resolved by a consensus algorithm that determines which transaction to keep based on the number of descendant transactions.

The recently proposed SPECTRE [119] and MeshCash [11] blockchains also use a DAG structure along with a protocol to reach consensus in the case of conflicts. Both are blockchains based on proof-of-work, which eliminates them from consideration for our use cases. SPECTRE's successor, PHANTOM [121]

requires strong network connectivity between honest nodes. HashGraph [6] does not rely on proof-of-work, but still requires strong network connectivity between members.

The DAG structure in the aforementioned blockchains is not designed to provide partition tolerance like our case, but rather to exploit available parallelism for increased throughput of transactions by only ordering transactions that are dependent. As such, these blockchains expect strong network connectivity and are therefore unsuitable in our use cases.

Aside from blockchains, there are a number of non-blockchain distributed systems related to ours. Bayou, a distributed storage system for low-power mobile devices with poor network connectivity is probably one of the closest works to ours [125]. Similarly, the COPS key-value store provides causal consistency in wide-area networks [80]. Both Bayou and COPS have ad-hoc merging protocols that require the application running on top of them to actively detect and resolve conflicts, unlike the transparent merging and precise semantics that Vegvisir provides.

In 2011, Mark Shapiro formalized the types of data structures that can be replicated across multiple hosts and updated concurrently and independently, while still providing strong eventual consistency [113]. These data structures, known as Conflict-free Replicated Data Types (CRDT), have been shown to include versions of registers, counters, sets, graphs, and maps [112]. These basic data types can be combined and composed to create more sophisticated data structures such as key-value stores [110] and JSON documents [70]. Applications include collaborative editing [82] and distributed databases [5].

Our blockchain uses a gossip-style protocol. Gossip protocols originated in the field of distributed databases [34], but has seen a resurgence coinciding with the proliferation of cloud computing [78, 27, 23] and more recently in blockchain protocols such as Bitcoin [71]. While a variant of the gossip protocol has been shown to work well in unreliable networks [125], most gossip protocols assume full network connectivity and can therefore not be directly applied in IoT environments with low connectivity.

4.5 Architecture

4.5.1 Design Requirements

The blockchain is essentially a log of records that are generally called *transactions* in the blockchain literature. The blockchain is maintained by a group of *users*. We would like the Vegvisir blockchain to have the following informal properties:

- *Tamperproof*: Once a transaction has been stored on the blockchain, it cannot be removed or modified, and neither can transactions that precede it in the blockchain.
- *Provenance*: If a user can read a transaction on the blockchain, then the user can read all transactions that precede it on the blockchain.
- *Authenticity*: Every transaction on the blockchain is identified by the user that created the transaction and placed it on the blockchain.

- *Transitivity*: If one user learns of a transaction on the blockchain, then eventually all users do.
- *Access Control*: There should exist control over which users are allowed to append which types of transactions to the blockchain.
- *Partition Tolerance*: The blockchain is available even when not all users can physically communicate with one another for some unspecified length of time.
- *Storage Efficiency*: IoT devices may have limited storage. They do not have to store all of the blockchain—some of it may be stored elsewhere.

These requirements, and in particular partition-tolerance, stipulate that the blockchain maintain a partial order of transactions. The transactions within a block are totally ordered, but a block may have multiple "parents." Nonetheless, Vegvisir will make an effort to reduce branching as much as possible. In particular, when a user appends a new transaction, all transactions known to the user must become ancestors of the transaction. Thus the Vegvisir blockchain maintains the causal history of all transactions.

4.5.2 Adversary Model

We assume that among the k closest network neighbors of a user (which may be malicious), at least one user correctly follows the Vegvisir protocol. The parameter k can be set according to need. Malicious peers want to change or remove blocks from the blockchain. Adversaries cannot forge signatures from other users, but they can remove blocks from their local version of the blockchain and they can choose not to propagate new blocks they receive.

4.5.3 Design Overview

Like most blockchains, Vegvisir consists of a series of interlinked blocks containing a block header and one or more transactions. Unlike most blockchains, each block can point to multiple other blocks as its predecessors. Thus blocks in Vegvisir form a DAG rather than a linear chain (see figure 4.1). There must be a single block, the genesis block, that is the ancestor of all other blocks.

The DAG structure of the chain, combined with CRDTs, is what makes Vegvisir partition-tolerant. If we were to require blocks to form a linear chain, we would have no choice but to either prevent blocks being added in all but one partition, or discard blocks when merging forks in the chain that have arisen due to network partitions. Preventing block from being added is unacceptable, but discarding violates tamperproofness and is therefore unacceptable as well. This is why we ruled out a linear chain.

The DAG encodes a partial ordering on transactions. When interpreting a DAG of transactions, we require that transactions that are not ordered with respect to one another in some sense commute. For this reason, we limit usage of Vegvisir to CRDT-based applications. The commutativity of CRDT operations removes the need for imposing a total order on transactions. Using CRDTs, any total ordering consistent with the partial ordering will produce the same interpretation on the state produced by the transactions. Below we will assume CRDT-based applications.

Vegvisir is a so-called *permissioned* blockchain and has a membership (for example, emergency first responders). It has an *owner* who generates and signs the genesis block. The genesis block contains a self-signed certificate of the owner,

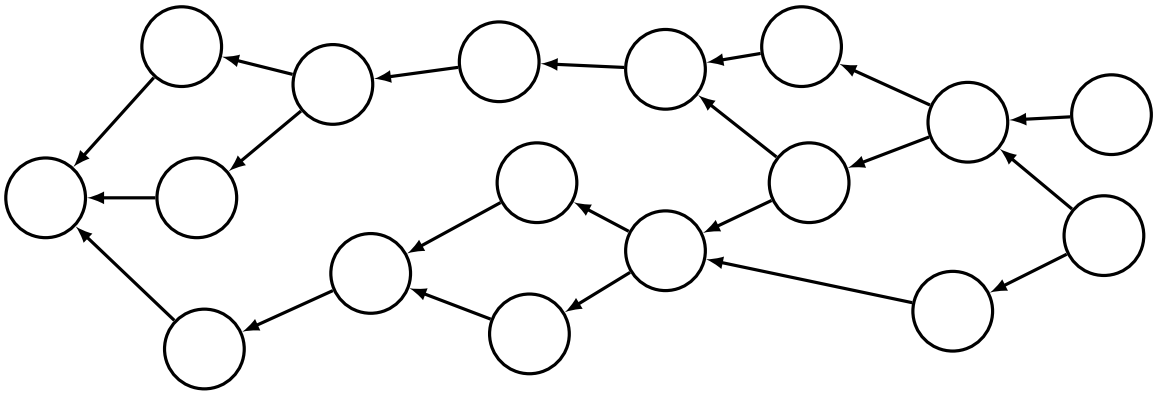


Figure 4.1: Sample DAG. Branches are reined in by making every known leaf a predecessor of your new block. As our applications are based on CRDTs, there is no need to determine a total order between the blocks.

who will act as a certificate authority (CA) on the blockchain. Each authorized user must have a certificate signed by the CA placed on the blockchain. Certificates specify the role of each user, and access control is determined based on those roles.

4.5.4 Blocks and Transactions

Each block is composed of block header, zero or more transactions, and a digital signature. The block header contains the user ID of the block creator, a timestamp, if possible a physical location, and a list of hashes of its parent blocks (see figure 4.2).

Transactions specify operations on CRDTs. For example, in our emergency first responder use case, a user might want to add an access request for a health record on the blockchain. Vegvisir could have an add-only set (which is a CRDT) of health record access requests. Call this CRDT $\mathcal{H}$. Then the user would add a transaction r, containing the request, to $\mathcal{H}$.

A transaction must specify the name of the CRDT, the type of operation to perform, and any arguments that operation requires. Transactions do not carry a digital signature—a transaction is implicitly signed by the block that contains the transaction. In Vegvisir, the creator of a block is the originator of all transactions in the block, so the block signature also establishes the authenticity and integrity of the transactions.

The set of valid users can also be thought of as a CRDT. Specifically, it is a *2P set*, which is a set representation composed of an *add set* A and a *remove set* R. When adding an element, it is added to A and when removing an element, it is added to R. The elements that are said to exist in the 2P set are $A \backslash R$. If the elements of A and R are public key certificates, then certificates can be added to A, while revocations amount to adding the same certificate to R. Every Vegvisir blockchain must have a 2P set of users, $\mathcal{U}$, and so it is implicitly created when a new blockchain is formed.

Other CRDTs, such as the add-only set $\mathcal{H}$ mentioned above, can be created as needed. Each new CRDT must have a unique name. To avoid collisions, names can be a randomly generated string of length n, where n is high enough that the probability of a naming collision is negligible. A collection of CRDTs is a CRDT itself. We will refer to the set of user-created CRDTs as Ω from now on.

4.5.5 Separation of Concerns

The software of each Vegvisir user has two main components. The first component is the blockchain. It maintains the local copy of the DAG, checks the validity of the blocks, and passes the transactions to the other component, the

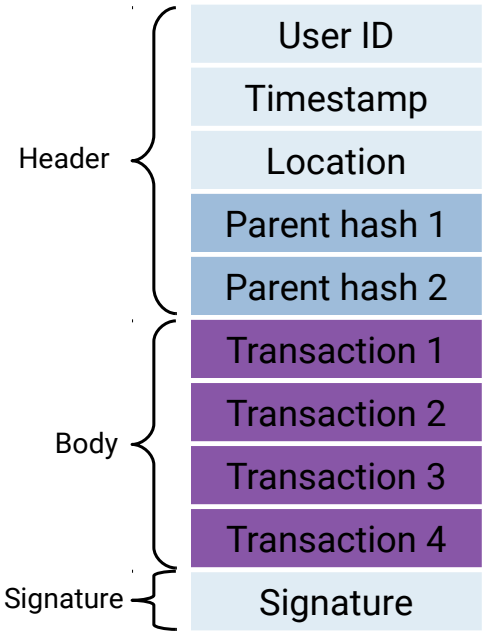


Figure 4.2: A layout of a block. The header contains a user ID, a timestamp, and if possible, a location. It also contains a variable number of parent hashes. The body consists of zero or more 'transactions'. Every block is signed by its creator, which is also the creator of all transactions in the block.

CRDT state machine (CSM). The only CRDT the blockchain component concerns itself with is $\mathcal{U}$. The following checks are performed to assess if a new block is valid:

- The user must be member of blockchain (specified by $\mathcal{U}$);
- Parent blocks must be in the blockchain already;
- The timestamp must be higher than the maximum of the timestamps in the parent blocks but lower than the current time at the user;
- The signature must be valid and match user ID.

The CSM in turn checks the validity of the transactions themselves and makes the appropriate updates to Ω and $\mathcal{U}$ once it has verified that the transaction satisfies the following:

- The CRDT must exist (i.e., it must be U, Ω, or an element of Ω);
- The specified operation must be valid for the CRDT;
- The argument to the operation must pass type checks (e.g. we cannot add an integer to a set of strings);
- The user must have permission to perform the operation.

When creating a CRDT, one must specify which roles can perform which actions. For example in the case of $\mathcal{H}$, it could be specified that only users with the role 'medic' can perform the add operation. Users' roles are specified in their public key certificates.

4.5.6 Public Key Certificates

A public key certificate contains the user ID, the public key of the user, the user's role, and a digital signature from the CA (the blockchain owner). When performing block validation, the user ID in the block header must match a user ID of one of the certificates in $\mathcal{U}$. Elements in the remove set of $\mathcal{U}$ act as certificate revocations. Similarly, when performing transaction validation, the CRDT indicates whether the user's role is permitted to perform the specified operation.

4.5.7 Opportunistic Reconciliation

Blocks are spread throughout the network via a protocol that resembles gossip. Periodically, a node picks a physical neighbor at random (if it has any). The initiator then asks the other node, the selected neighbor, for its *frontier set*. The frontier set is the set of blocks on the DAG that have no successors (sources, given that blocks in the DAG point to their parent blocks). If the neighbor's frontier set is identical to the initiator's, then their blockchains are identical too and the process stops. If, however, the frontier set contains blocks unknown to the initiator, it adds the frontier set to its own replica of the DAG. That operation will fail if the DAG does not contain all parents of all blocks in the frontier set. In that case, the initiator requests to see the level 2 frontier set, which is the frontier set plus the set of all parent nodes. In general, a level N frontier set is defined as the union of the level $N-1$ frontier set and the parents of all blocks in the $N-1$ frontier set. The base case of this recursive definition is the level 1 frontier set, which is the frontier set described above (see figure 4.3). The initiator continues to ask for higher levels of the frontier set until it is able to bridge the gap between

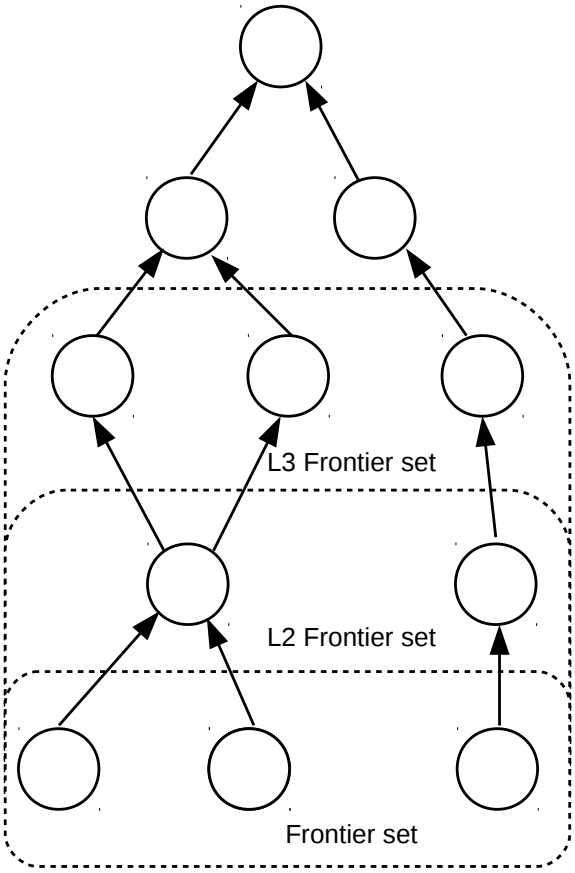


Figure 4.3: The (level 1) frontier set is the set of blocks without successors. The level 2 frontier set contains the level 1 frontier set plus their parents. More generally, the level n frontier set is the union of the level $n-1$ frontier set and the parents of its blocks.

its blockchain and the neighbor's. That must happen eventually assuming they have the same genesis block (which is the unique sink of the DAG and identifies the Vegvisir blockchain).

4.5.8 Persistence through Proof-of-Witness

Persistence and immutability are primary advantages of blockchains. Since malicious nodes may attempt to remove newly added blocks, we cannot be confi-

Algorithm 1 DAG Reconciliation Pseudocode

```
1:  procedure RECONCILIATEDAGS(S)                          ▷ S: local DAG
2:      B ← getRandomNeighbor()
3:      if B is not empty then
4:          n ← 1
5:          S_{B,n} ← getNthFrontierSet(B, n)
6:          if parents(S_{B,n}) ⊆ S then
7:              S ← merge(S_{B,n}, S)
8:          else
9:              n ← n + 1
10:             goto 5
11:         end if
12:     end if
13: end procedure
```

dent that a block will persist once it is added by one user of the blockchain. In particular, an application may not be able to take action until it has some guarantee that a particular transaction and the transactions that causally precede it are persistent. To solve this problem, an application may require confirmation from users in some quorum that they have a copy of the block. The choice of quorum is up to the application. Because the Vegvisir blockchain is a DAG rather than a linear chain, there is no requirement that quorums overlap.

For example, if a user requests access for a health record by adding a transaction to $\mathcal{H}$, an application may require that k additional nearby users have stored the block containing the transaction before counting it as a persistent part of the blockchain. One way to obtain the desired effect is as follows: A user may indicate that it has stored a block by adding an ancestor block to the blockchain, signed by that user. Once a block has ancestor blocks signed by at least k different nearby users, the block may be considered persistent by the application. These blocks need not contain any transactions. Their sole purpose is to signal that a user has a copy of the ancestor blocks. We say that a block has a *proof-of-*

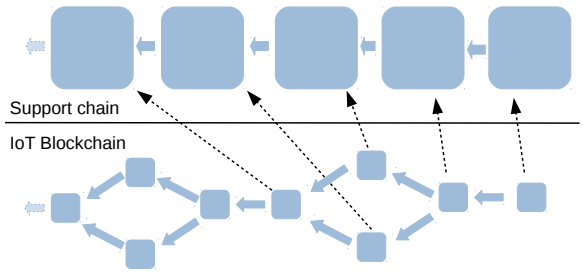


Figure 4.4: The IoT blockchain has periodic access to a support blockchain.

witness once it has reached this condition. Note that a proof-of-witness does not only apply to the block itself, but also to each of its ancestor blocks.

4.5.9 Support Blockchain

IoT devices may be constrained by the amount of storage they have for the Vegvisir blockchain. We allow such devices to offload parts of their DAG to a more traditional blockchain that we call the *support blockchain* (see figure 4.4). This would be applicable to environments where the low-power, battery-constrained IoT devices that make up the partition-tolerant blockchain have occasional access to higher-powered servers. The higher-powered servers can function as superpeers, taking blocks from the partition-tolerant blockchain and placing them on the support blockchain, which operates between the superpeers as well as in the cloud. Once a block is placed on the support blockchain, the IoT device can drop the block. Typically, IoT devices would only do so when running low on storage, and would only offload their oldest blocks on the blockchain.

As superpeers get new blocks, they in turn add new blocks to the support blockchains. The body of a block on the support blockchain is a Vegvisir block.

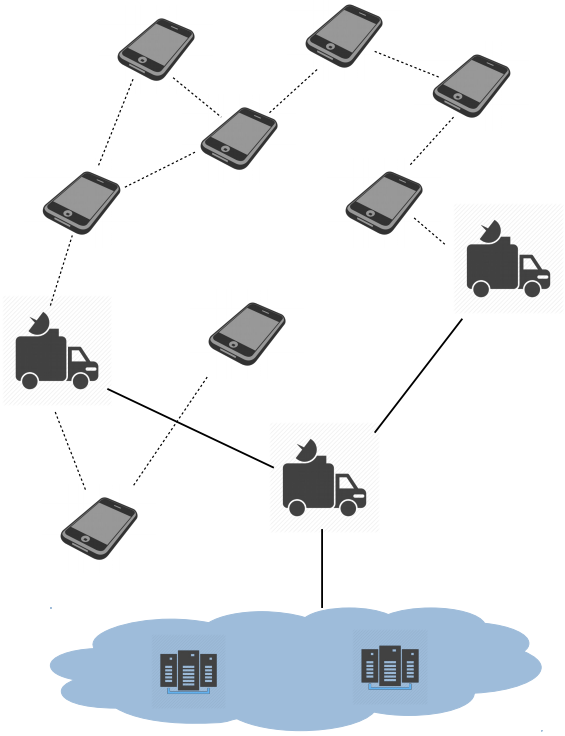


Figure 4.5: The network can consist of both battery-constrained IoT devices (depicted as smartphones) as well as relatively high-powered deployable servers (depicted as trucks) that may be connected to the rest of the Internet. The high-powered nodes relay blocks to the support.

Support blocks must be added in a way that preserves the topological order of the Vegvisir DAG.

4.6 Implementation

We have an Android implementation of Vegvisir under construction. The Android prototype is designed specifically for the emergency-first responder scenario. As such, it allows users to place requests for health records on the blockchain. It uses Bluetooth and Google Nearby (which uses a combination of Bluetooth and WiFi Direct) to communicate opportunistically with anyone in

its neighborhood. Additionally, the implementation has to have a mechanism to deliver the health records to the requester once the request is stored securely on the blockchain.

The simplest way would be to have the user present a proof-of-witness that their request has been placed on the blockchain to a centralized database server. But a key assumption of Vegvisir is that devices operate in an environment with unreliable access to each other and the public cloud, so the device might not be able to connect to such a database for extended periods of time. An alternative could be to have each user carry an encrypted version of the database in secondary storage. The key would be kept in a Trusted Execution Environment (TEE) and only through a certifiably correct program can a health record be decrypted and made available to the user and only once that program has determined that the request is on the blockchain and has obtained the proof-of-witness.

4.7 Conclusion and Future Work

Vegvisir is a prime example of how the IoT can be leveraged to enhance privacy instead of undermine it. By forming a highly partition-tolerant distributed system that can implement any CRDT, it enables a tamperproof log in environments with limited to no internet connectivity but with limited access to a number of IoT devices. A tamperproof log can be used as a building block for secure and privacy aware systems, thus leveraging the IoT for security and privacy in a novel way.

While we could implement Vegvisir without using CRDTs and design ad hoc

merging protocols for our use cases, we believe CRDTs provide a sweet spot in our design space: they allow us to operate in a partitionable environment while still giving precise semantics on shared data structures. The alternative of providing linearizability would have led to lack of liveness, while the alternative of ad hoc merging (as in [125, 80]) lacks well-defined semantics.

There is nonetheless room for improvement in Vegvisir. The opportunistic reconciliation method, while considerably more efficient than exchanging entire DAGs, still incurs a significant communication overhead. More efficient DAG reconciliation algorithms could make blocks propagate faster through the network while using less bandwidth.

Providing partition-tolerance is vital if blockchains are to be adapted to IoT environments. IoT devices operate with strict constraints on power and often limited network connectivity. Vegvisir extends partition tolerance to blockchains, although at the cost of limiting the classes of applications that can be implemented with CRDTs. While that prohibits applications that need a unique total ordering on conflicting transactions, like cryptocurrencies where double spending must be prevented, it can nonetheless be used to implement a persistent, tamperproof distributed ledger suitable as a building block in many useful applications. We will review two potential avenues of research within the fields of digital agriculture and maritime forensics.

4.7.1 Digital Agriculture

The blockchain is a promising technology that could bring transparency and accountability to the food supply chain. There are many participants in the

food supply chain, including farmers, brokers, packers, traders, distributors, food processors, retailers, regulators, and ultimately, consumers. Today, record keeping is mostly on paper and various centralized databases, while many negotiations are purely verbal. This is prone to mistakes and simplifies fraud. Farmers do not know how much profit is made on the food they supply, and consumers cannot easily track where their food comes from. Blockchains could potentially create a shared and tamperproof data repository in which all information is readily shared, available, and auditable.

Blockchains could make it straightforward for a consumer to check the source of a food product. In the case of a meat product, information of interest might include the animal's date of birth, place of origin, vaccinations, and use of antibiotics. Food safety is a related application. For example, Walmart (which sells 20% of all food in the U.S.), IBM, and Tsinghua University are looking into using the Hyperledger blockchain [18] for food supply chain traceability and authenticity. Today, if a pathogen is found in a food product, it takes several days to weeks to trace it back to the supplier. Using a blockchain, Walmart hopes to reduce this to seconds, potentially saving lives. From the farmer's perspective, blockchains could make it easier to find consumers for their products, potentially reducing waste from either unused land (if too little is produced) or from overproduction. Farmers could check to see what retailers sell their products for, so they may negotiate a better price for their produce.

Ideally, record keeping with the blockchain would reach all the way to the farm and to the distribution channels. Tagging of animals, pallets, shipping containers, and so on with RFIDs or related technology enable tracking. But farms and distribution centers, let alone sensors, autonomous vehicles, and drones

operating out in the field, have intermittent if any connection to the Internet, and must rely on a system consisting of small fixed and mobile IoT devices for sensing. While such systems can be used to create a sensor network, there is no integrated blockchain that can securely store the history of sensed data. Again, Vegvisir allows deploying a low-energy tamperproof log in this environment.

4.7.2 Marine Forensics

According to the National Oceanic and Atmospheric Administration, the U.S. maritime transportation infrastructure was responsible for carrying $1.5 trillion in cargo in 2017 [96]. An individual ship can hold a fortune in terms of technology, crew, and data. The loss of a single ship could be devastating to a company's future.

The information contained within ship systems can present valuable insight that can prevent future mishaps and save lives [59]. Ships can sink in solitude within a matter of minutes making it dangerous and impractical to attempt data collection under these circumstances. Moreover, retrieval of data from sunken vessels can take months to obtain and physical information from these events are sometimes unavailable due to contamination, water depth, or miscellaneous damage to instrumentation.

Vegvisir's structure can allow for data collection on trade ships during capsizing events. Distress signals already sent out during ship emergencies could trigger the construction of an ad hoc network used by select systems on the vessel. IoT devices on lifeboats could autonomously join the network at the time of their inflation.

As new nodes emerge on the network, Vegvisir's opportunistic gossip protocol can commence. Vegvisir creates a low-power consumption structure for its blockchain. This ensures that even as IoT devices acquire different information from various sources, minimal energy is spent on blockchain reconciliation. In the event of a submersion, the lifeboat nodes would still be able to gossip amongst themselves.

Due to the nature of proprietary information in the maritime industry, Vegvisir security guarantees would be amenable in these environments. First, the secure membership protocol prevents non-verified members from contributing events. Additionally, all blocks are signed by the contributing member. Therefore, blockchain data consists only of data from approved sources regardless of network activity. Second, Vegvisir allows for full encryption of contents within the blockchain. These policies in conjunction serve as safeguards for company proprietary information sent over its ad hoc network while maintaining its properties of possessing tamperproof logs and being energy-efficient.

4.8 Appendix

The central privacy property of Vegvisir is that it allows enforcement of an accountability policy. No node in the blockchain network will be given access to a sensitive record unless their request is securely entered onto the blockchain. We say a block B has been securely entered onto the blockchain if eventually all honest nodes learn block B. An honest node is a member of the blockchain, a replica in the distributed state machine, that honestly and correctly follows the protocol. The Vegvisir protocol has been designed to fulfill the following three

informal properties:

1. No honest node will throw away a valid block
2. If an honest node learns a block, eventually all honest nodes will learn that block
3. At most $k - 1$ nodes are not honest

Strictly speaking, (2) implies (1), but it is included for conceptual clarity. If these assumptions hold, then Vegvisir must be tamperproof, i.e. if a quorum of k nodes has learned of a block B, then eventually all honest nodes will learn of B. To see that, note that if B is learned by k different nodes (i.e. reaches a quorum), at least one of those nodes must be honest by assumption (3). Therefore, by assumption (2), eventually all honest nodes will learn of B. We conclude that once a block reaches a quorum, its contents are tamperproof.

CHAPTER 5

CONCLUSION

5.1 Summary

In this dissertation, I examined the design and analysis of privacy mechanisms for the Internet of Things. We began by considering internet advertising. Internet advertising is a particularly interesting case because it is simultaneously the economic driving force undermining privacy in the IoT as well a mechanism for data collection in itself. I examined not only how internet advertising works but also what happens to the information collected. This was followed by a review of the first type of privacy mechanisms we encountered in the dissertation: ad blockers.

I then moved on to some of the most sensitive information collected by the IoT: location information. We showed that most research up to date focused on protecting the identity or physical location of the user, while in many cases, it is the semantics the location that needs to be protected. I developed tools to analyze semantic location privacy based on Bayesian decision theory and Differential Privacy. I showed how location-based social media data can be used to evaluate location privacy mechanisms and I developed a theoretical framework for analysis based on spatial point processes. Finally, I presented a location privacy mechanism designed specifically to provide semantic location privacy without sacrificing too much quality of service. I analyzed said mechanism using all the tools developed in the chapter, including a combination of the Bayesian and Differential Privacy approaches as well as using both empirical and theoretical approaches.

Lastly, I showed how the Internet of Things can be used to enhance privacy instead of undermine it. Using a blockchain of our own design, Vegvisir, I show how to leverage the IoT to ensure the security and privacy of sensitive records such as medical records during disaster response. I examined the design of Vegvisir and provided an analysis of its security and privacy properties.

5.2 Future Directions

The design and analysis of privacy mechanisms for the IoT is still in its infancy. The IoT presents previously unencountered privacy problems, requiring novel solutions and engineering analysis. While we have developed a toolbox for the analysis of location privacy mechanisms in the IoT, there is still more work to do. A part of that work revolves around exploring alternative spatial point process models to model semantic locations, such as Matern, Strauss, and Thomas processes could form an interesting line of research. And while we showed how some of the theoretical properties of differential privacy extend to semantic geo-indistinguishability, there is still more work to be done on that front. Furthermore, the randomized response mechanism we presented has room for improvement. Most pressingly, the partition created using the balanced graph partitioning algorithm is based on heuristics, as there is no know tractable optimal algorithm. It could potentially be improved by combining the heuristic graph partitioning with a local search algorithm that would use the graph partitioning as a starting point and then incrementally improve it.

There are also plenty of avenues for further research on Vegvisir. As a partition-tolerant blockchain, it could potentially have considerable impact be-

yond the field of privacy. Any scenario where low-powered devices with limited network connectivity need to form a shared data repository that is resilient, tamperproof, and uses a distributed trust model could potentially benefit from Vegvisir. Digital agriculture and marine forensics are two prime examples of areas that could have uses for Vegvisir. Vegvisir furthermore needs more thorough systems evaluation to determine its resource consumption, as well as the conditions under which proof-of-witness can be obtained in a reasonable amount of time.

As the Internet of Things is projected to grow exponentially in the next decade, there will undoubtedly be countless other problems to explore in the coming years. In this thesis, we attempted to provide a principled starting point for some of the most pressing privacy problems in the IoT today. Hopefully, such research will continue to gain traction as the IoT and its associated privacy problems continue to grow.

BIBLIOGRAPHY

[1] Baltimore & Ohio R. Co. v. United States, 1923. 261 U.S. 592, Supreme Court of the United States.

[2] Carpenter v. United States, 2018. 585 U.S. ___, Supreme Court of the United States.

[3] Jon Alexander, Tom Crompton, and Guy Shrubsole. Think of Me as Evil? Opening the Ethical Debates in Advertising. Technical report, Public Interest Research Center, October 2011.

[4] Miguel E. Andrs, Nicols E. Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Geo-indistinguishability: Differential Privacy for Location-based Systems. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, CCS '13, pages 901–914, New York, NY, USA, 2013. ACM.

[5] Cihan B. Getting Started with Active-Active Geo-Distribution for Redis Applications with CRDTs (Conflict-free Replicated Data Types), October 2017.

[6] Leemon Baird. The Swirlds hashgraph consensus algorithm: Fair, fast, Byzantine fault tolerance. Technical report, Swirlds Tech Report SWIRLDS-TR-2016-01, available online, http://www. swirlds. com/developer-resources/whitepapers, 2016.

[7] Bhuvan Bamba, Ling Liu, Peter Pesti, and Ting Wang. Supporting Anonymous Location Queries in Mobile Environments with Privacygrid. In *Proceedings of the 17th International Conference on World Wide Web*, WWW '08, pages 237–246, New York, NY, USA, 2008. ACM.

[8] Jane Bambauer, Krishnamurty Muralidhar, and Rathindra Sarathy. Fool's Gold: An Illustrated Critique of Differential Privacy. *Vanderbilt Journal of Entertainment & Technology Law*, 16(4), 2014.

[9] Abigail Beall. Bitcoin mining uses more energy than Ecuador but theres a fix. *New Scientist*, October 2017.

[10] Jeremy Bentham. *A Fragment on Government*. History of Economic Thought Books. McMaster University Archive for the History of Economic Thought, London, 1776.

[11] Iddo Bentov, Pavel Hubcek, Tal Moran, and Asaf Nadler. Tortoise and Hares Consensus: the Meshcash Framework for Incentive-Compatible, Scalable Cryptocurrencies. *IACR Cryptology ePrint Archive*, 2017:300, 2017.

[12] A. R. Beresford and F. Stajano. Mix zones: user privacy in location-aware services. In *IEEE Annual Conference on Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second*, pages 127–131, March 2004.

[13] Francine Berman and Vinton G. Cerf. Social and Ethical Behavior in the Internet of Things. *Commun. ACM*, 60(2):6–7, January 2017.

[14] Eleanor Birrell and Fred B. Schneider. A Reactive Approach for Use-Based Privacy. Technical report, November 2017. http://hdl.handle.net/1813/54843.

[15] Avrim Blum, Katrina Ligett, and Aaron Roth. A Learning Theory Approach to Noninteractive Database Privacy. *J. ACM*, 60(2):12:1–12:25, April 2013.

[16] Stefano Bocconi, Alessandro Bozzon, Achilleas Psyllidis, Christiaan Titos Bolivar, and Geert-Jan Houben. Social Glass: A Platform for Urban Analytics and Decision-making Through Heterogeneous Social Data. In *Proceedings of the 24th International Conference on World Wide Web Companion*, WWW '15 Companion, pages 175–178, Republic and Canton of Geneva, Switzerland, 2015. International World Wide Web Conferences Steering Committee.

[17] Mark Bun and Thomas Steinke. Concentrated Differential Privacy: Simplifications, Extensions, and Lower Bounds. In *Theory of Cryptography*, Lecture Notes in Computer Science, pages 635–658. Springer, Berlin, Heidelberg, November 2016.

[18] Christian Cachin. Architecture of the Hyperledger blockchain fabric. In *Workshop on Distributed Cryptocurrencies and Consensus Ledgers*, 2016.

[19] Xin Cao, Gao Cong, and Christian S. Jensen. Mining Significant Semantic Locations from GPS Data. *Proc. VLDB Endow.*, 3(1-2):1009–1020, September 2010.

[20] Fred Cate. Principles for protecting privacy. *Cato Journal*, 22(1):33–58, 2002.

[21] Fred Cate, Peter Cullen, and Viktor Mayer-Schonberger. *Data Protection Princinples for the 21st Century*. Number 23 in Books by Maurer Faculty. 2013.

[22] U. S. Census Bureau. American FactFinder, 2018. Published: (2015, June, 10). [Online]. Available: http://factfinder.census.gov.

[23] R. Chandra, V. Ramasubramanian, and K. Birman. Anonymous Gossip: improving multicast reliability in mobile ad-hoc networks. In *Proceedings 21st International Conference on Distributed Computing Systems*, pages 275–283, April 2001.

[24] Yanzhe Che, Kevin Chiew, Xiaoyan Hong, and Qinming He. SALS: Semantics-aware Location Sharing Based on Cloaking Zone in Mobile Social Networks. In *Proceedings of the First ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems*, MobiGIS '12, pages 49–56, New York, NY, USA, 2012. ACM.

[25] Zhiyuan Cheng, James Caverlee, Kyumin Lee, and Daniel Z Sui. Exploring Millions of Footprints in Location Sharing Services. In *Proceeding of the 5th International AAAI Conference on Weblogs and Social Media (ICWSM)'11*, Barcelona, Spain, 2011.

[26] Konstantinos Christidis and Michael Devetsikiotis. Blockchains and smart contracts for the internet of things. *IEEE Access*, 4:2292–2303, 2016.

[27] L. Chuat, P. Szalachowski, A. Perrig, B. Laurie, and E. Messeri. Efficient gossip protocols for verifying the consistency of Certificate logs. In *2015 IEEE Conference on Communications and Network Security (CNS)*, pages 415–423, September 2015.

[28] Anton Churyumov. Byteball: A Decentralized System for Storage and Transfer of Value. Technical report, 2015. https://byteball.org/Byteball.pdf.

[29] Federal Trade Commission. Data Brokers: A Call For Transparency and Accountability. Technical report, May 2014.

[30] Maria Luisa Damiani, Elisa Bertino, and Claudio Silvestri. Protecting Location Privacy Against Spatial Inferences: The PROBE Approach. In *Proceedings of the 2Nd SIGSPATIAL ACM GIS 2009 International Workshop on Security and Privacy in GIS and LBS*, SPRINGL '09, pages 32–41, New York, NY, USA, 2009. ACM.

[31] Maria Luisa Damiani, Elisa Bertino, and Claudio Silvestri. Protecting Location Privacy Against Spatial Inferences: The PROBE Approach. In *Proceedings of the 2Nd SIGSPATIAL ACM GIS 2009 International Workshop on Security and Privacy in GIS and LBS*, SPRINGL '09, pages 32–41, New York, NY, USA, 2009. ACM.

[32] Quynh H. Dang. Secure Hash Standard (SHS). NIST Pubs 180-4, National Institute of Standards and Technology, March 2012.

[33] Jacob Davidson. Here's How Apple Could Change the Web Forever. *Time*, September 2015.

[34] Alan Demers, Dan Greene, Carl Hauser, Wes Irish, John Larson, Scott Shenker, Howard Sturgis, Dan Swinehart, and Doug Terry. Epidemic Algorithms for Replicated Database Maintenance. In *Proceedings of the Sixth Annual ACM Symposium on Principles of Distributed Computing*, PODC '87, pages 1–12, New York, NY, USA, 1987. ACM.

[35] Peter Diggle. *Statistical analysis of spatial and spatio-temporal point patterns*. CRC Press, Taylor & Francis Group, Boca Raton, third edition. edition, 2014. http://newcatalog.library.cornell.edu/catalog/8431181.

[36] Tewfik Doumi, Mike F Dolan, Said Tatesh, Alessio Casati, George Tsirtsis, Kiran Anchan, and Dino Flore. LTE for public safety networks. *IEEE Communications Magazine*, 51(2):106–112, 2013.

[37] Leonard Downie, Jr. and Michael Schudson. The Reconstruction of American Journalism. Technical report, Columbia University, October 2009.

[38] Cynthia Dwork. Differential Privacy. In *Automata, Languages and Programming*, Lecture Notes in Computer Science, pages 1–12. Springer, Berlin, Heidelberg, July 2006.

[39] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating Noise to Sensitivity in Private Data Analysis. In *Theory of Cryptography*, Lecture Notes in Computer Science, pages 265–284. Springer, Berlin, Heidelberg, March 2006.

[40] Morris J. Dworkin. SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions. NIST Pubs 202, National Institute of Standards and Technology, August 2015.

[41] AllBusiness Editors. Web Advertising and CPM: A Quick Guide for Small Businesses. *AllBusiness.com*, September 2011.

[42] lfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, CCS '14, pages 1054–1067, New York, NY, USA, 2014. ACM.

[43] Gabor Fodor, Stefan Parkvall, Stefano Sorrentino, Pontus Wallentin, Qianxi Lu, and Nadia Brahmi. Device-to-device communications for national security and public safety. *IEEE Access*, 2:1510–1520, 2014.

[44] Electronic Frontier Foundation. Turner CEO Says Bathroom Breaks are Theft. *EFFector*, 15(15), May 2002.

[45] Electronic Frontier Foundation. Privacy Badger, April 2013. [Software]. https://www.eff.org/privacybadger.

[46] OpenStreetMap Foundation. OpenStreetMap, 2018. Published: (2015, June, 10). [Online]. Accessed: http://www.openstreetmap.org.

[47] Inc. Foursquare Labs. Foursquare, 2018. Published: (2015, June, 2nd) [Online]. Available: https://foursquare.com/.

[48] Foursquare Labs, Inc. About Foursquare, 2018. Published: (2015, June, 12). [Online]. Accessed: https://foursquare.com/about.

[49] William Frankena. *Ethics*. Prentice Hall, Upper Saddle River, 2 edition, 1973.

[50] Brandon Gaille. 26 Great Foursquare Demographics, January 2015. Published: (2015, June, 12). [Online]. Available: http://brandongaille.com/26-great-foursquare-demographics/.

[51] Alan E. Gelfand, Peter Diggle, Peter Guttorp, and Montserrat Fuentes. *Handbook of Spatial Statistics*. CRC Press, March 2010.

[52] Samuel Gibbs. Apple removes adblockers that work on Facebook and other third-party apps. *The Guardian*, October 2015.

[53] Seth Gilbert and Nancy Lynch. Brewer's Conjecture and the Feasibility

of Consistent, Available, Partition-tolerant Web Services. *SIGACT News*, 33(2):51–59, June 2002.

[54] Philippe Golle and Kurt Partridge. On the Anonymity of Home/Work Location Pairs. In Hideyuki Tokuda, Michael Beigl, Adrian Friday, A. J. Bernheim Brush, and Yoshito Tobe, editors, *Pervasive Computing*, number 5538 in Lecture Notes in Computer Science, pages 390–397. Springer Berlin Heidelberg, 2009.

[55] Marco Gruteser and Dirk Grunwald. Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking. In *Proceedings of the 1st International Conference on Mobile Systems, Applications and Services*, MobiSys '03, pages 31–42, New York, NY, USA, 2003. ACM.

[56] Drew Harwell. Whirlpools Internet of Things problem: No one really wants a smart washing machine. *Washington Post*, October 2014.

[57] Robert Hof. Mobile Ads Will Smash $100 Billion Mark Worldwide In 2016. *Forbes*, April 2015.

[58] Sean Hollister. Facebook's ads have been defeated (again) by Adblock Plus work-around. *CNET*, August 2016.

[59] Joshua Howgego. Peril on the sea: Why are so many megaships sinking? *New Scientist*, January 2018.

[60] Joel Hruska. Mozilla Firefox kills Flash by default, security chief calls for Adobe to issue an end-of-life date. *ExtremeTech*, July 2015.

[61] Chi-Min Huang, J.J.-C. Ying, V.S. Tseng, and Zhi-Hua Zhou. Location semantics prediction for living analytics by mining smartphone data. In *2014 International Conference on Data Science and Advanced Analytics (DSAA)*, pages 527–533, October 2014.

[62] IAB. Rothenberg Says Ad Blocking Is a War against Diversity and Freedom of Expression, January 2016. Accessed: https://www.iab.com/news/rothenberg-says-ad-blocking-is-a-war-against-diversity-and-freedom-of-expression.

[63] Apple Inc. What's New in iOS 9.0, June 2017. Accessed: https://developer.apple.com/library/archive/releasenotes /General/WhatsNewIniOS/Articles/iOS9.html.

[64] Legal Information Institute. Contract Implied in Fact, October 2009. [Encyclopedia article]. Accessed: https://www.law.cornell.edu/wex/contract_implied_in_fact.

[65] Jasper Jackson. Three network to run 24-hour adblocking trial. *The Guardian*, May 2016.

[66] Immanuel Kant. *Groundwork for the Metaphysics of Morals*. 1785.

[67] K. Karlsson and S. B. Wicker. The Effect of Location Granularity on Semantic Location Inferences. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 2197–2204, January 2016.

[68] G. Karypis and V. Kumar. A Fast and High Quality Multilevel Scheme for Partitioning Irregular Graphs. *SIAM Journal on Scientific Computing*, 20(1):359–392, January 1998.

[69] H. Kido, Y. Yanagisawa, and T. Satoh. An anonymous communication technique using dummies for location-based services. In *International Conference on Pervasive Services, 2005. ICPS '05. Proceedings*, pages 88–97, July 2005.

[70] M. Kleppmann and A. R. Beresford. A Conflict-Free Replicated JSON Datatype. *IEEE Transactions on Parallel and Distributed Systems*, 28(10):2733–2746, October 2017.

[71] Philip Koshy, Diana Koshy, and Patrick McDaniel. An Analysis of Anonymity in Bitcoin Using P2p Network Traffic. In *Financial Cryptography and Data Security*, Lecture Notes in Computer Science, pages 469–485. Springer, Berlin, Heidelberg, March 2014.

[72] John Krumm. Inference Attacks on Location Tracks. In Anthony LaMarca, Marc Langheinrich, and Khai N. Truong, editors, *Pervasive Computing*, number 4480 in Lecture Notes in Computer Science, pages 127–143. Springer Berlin Heidelberg, 2007.

[73] John Krumm. A Survey of Computational Location Privacy. *Personal Ubiquitous Comput.*, 13(6):391–399, August 2009.

[74] Ciphort Labs. The Rise of Malvertising. Technical report, 2015. http://go.cyphort.com/Malvertising-Report-15-Page.html.

[75] Byoungyoung Lee, Jinoh Oh, Hwanjo Yu, and Jong Kim. Protecting Location Privacy Using Location Semantics. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, pages 1289–1297, New York, NY, USA, 2011. ACM.

[76] N. Li, T. Li, and S. Venkatasubramanian. t-Closeness: Privacy Beyond k-Anonymity and l-Diversity. In *2007 IEEE 23rd International Conference on Data Engineering*, pages 106–115, April 2007.

[77] Yanhua Li, M. Steiner, Limin Wang, Zhi-Li Zhang, and Jie Bao. Exploring venue popularity in foursquare. In *2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 205–210, April 2013.

[78] JongBeom Lim, Kwang-Sik Chung, HwaMin Lee, Kangbin Yim, and Heonchang Yu. Byzantine-resilient dual gossip membership management in clouds. *Soft Computing*, pages 1–12, March 2017.

[79] Juhong Liu, O. Wolfson, and Huabei Yin. Extracting Semantic Location from Outdoor Positioning Systems. In *7th International Conference on Mobile Data Management, 2006. MDM 2006*, pages 73–73, May 2006.

[80] Wyatt Lloyd, Michael J. Freedman, Michael Kaminsky, and David G. Andersen. Don't Settle for Eventual: Scalable Causal Consistency for Wide-area Storage with COPS. In *Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles*, SOSP '11, pages 401–416, New York, NY, USA, 2011. ACM.

[81] Xuelian Long, Lei Jin, and J. Joshi. Understanding venue popularity in Foursquare. In *2013 9th International Conference Conference on Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom)*, pages 409–418, October 2013.

[82] Xiao Lv, Fazhi He, Weiwei Cai, and Yuan Cheng. A string-wise CRDT algorithm for smart and large-scale collaborative editing systems. *Advanced Engineering Informatics*, 33:397–409, August 2017.

[83] Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, and Muthuramakrishnan Venkitasubramaniam. L-diversity: Privacy Beyond K-anonymity. *ACM Trans. Knowl. Discov. Data*, 1(1), March 2007.

[84] Alasdaire MacIntyre. *After Virtue: A Study in Moral Theory*. University of Notre Dame Press, South Bend, 3 edition, 2007.

[85] Jack Marshall. Facebook Will Force Advertising on Ad-Blocking Users. *Wall Street Journal*, August 2016.

[86] Robert W. McChesnay and Victor Pickard, editors. *Will the Last Reporter Please Turn Out the Lights: The Collapse of Journalism and What Can Be Done to Fix It*. New Press, New York, 2011.

[87] Alfred J. Menezes, Jonathan Katz, Paul C. van Oorschot, and Scott A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, Boca Raton, 1 edition, October 1996.

[88] Andrew Meola. The US government is pouring money into the Internet of Things. *Business Insider*, May 2016.

[89] Claire Cain Miller. When Algorithms Discriminate. *The New York Times*, December 2017.

[90] I. Mironov. Rnyi Differential Privacy. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pages 263–275, August 2017.

[91] Robert L. Mitchell. Ad blockers: A solution or a problem? *Computerworld*, January 2014.

[92] Michael Mitzenmacher. A Brief History of Generative Models for Power Law and Lognormal Distributions. *Internet Mathematics*, 1(2):226–251, January 2004.

[93] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. Technical report, 2008.

[94] Nic Newman, David A. L. Levy, and Rasmus Kleis Nielsen. Reuters Institute Digital News Report 2015. *SSRN Electronic Journal*, 2015.

[95] Rishab Nithyanand, Sheharbano Khattak, Mobin Javed, and Narseo Vallina-Rodriguez. Adblocking and Counter Blocking: A Slice of the Arms Race. USENIX, August 2016.

[96] NOAA. How important is the ocean to our economy?, 2018. https://oceanservice.noaa.gov/facts/oceaneconomy.html.

[97] A. Noulas and C. Mascolo. Exploiting Foursquare and Cellular Data to Infer User Activity in Urban Environments. In *2013 IEEE 14th International*

Conference on Mobile Data Management (MDM), volume 1, pages 167–176, June 2013.

[98] Karl J O'Dwyer and David Malone. Bitcoin mining and its energy footprint. In *25th IET Irish Signals & Systems Conference 2014 and 2014 China-Ireland International Conference on Information and Communications Technologies (ISSC 2014/CIICT 2014)*. IET, 2014.

[99] Cathy O'Niel. *Weapons of Math Destruction*. Crown Books, 2016.

[100] Lara O'Reilly. This ad blocking company has the potential to tear a hole right through the mobile web and it has the support of carriers. *Business Insider*, May 2015.

[101] PageFair. Adblocking Goes Mobile. Technical report, November 2016.

[102] PageFair and Adobe. The 2015 Ad Blocking Report. Technical report, August 2015.

[103] Derek Parfit. *On What Matter*, volume 1. Oxford University Press, 2011.

[104] Madison Park and Michael Guy. Hurricane season is finally ending. *CNN*, November 2017.

[105] Avram Piltch. Why Using an Ad Blocker Is Stealing. *Tom's Guide*, May 2015.

[106] Sergui Popov. The Tangle. Technical report, April 2018. Accessed: http://iota.org/IOTA_Whitepaper.pdf.

[107] Dan Primack. Dear Apple: I may rob your store. *Fortune*, September 2015.

[108] D. Quercia and D. Saez. Mining Urban Deprivation from Foursquare: Implicit Crowdsourcing of City Land Use. *IEEE Pervasive Computing*, 13(2):30–36, April 2014.

[109] Randall Rothenberg. Ad Blocking: The Unnecessary Internet Apocalypse, September 2015. Accessed: http://adage.com/article/digitalnext/ad-blocking-unnecessary-internet-apocalypse/300470/.

[110] Salvatore Sanfilippo. Redis, 2009. [Software]. https://redis.io/.

[111] Thomas M. Scanlon. *What We Owe Each Other*. Belknap Press, Cambridge, MA, 2000.

[112] Marc Shapiro, Nuno Preguia, Carlos Baquero, and Marek Zawirski. A comprehensive study of Convergent and Commutative Replicated Data Types. report, Inria Centre Paris-Rocquencourt ; INRIA, January 2011.

[113] Marc Shapiro, Nuno Preguia, Carlos Baquero, and Marek Zawirski. Conflict-Free Replicated Data Types. In *Stabilization, Safety, and Security of Distributed Systems*, Lecture Notes in Computer Science, pages 386–400. Springer, Berlin, Heidelberg, October 2011.

[114] Clay Shirky. Newspapers and Thinking the Unthinkable. *Risk Management; New York*, 56(3):24–26,28–29,3, May 2009.

[115] R. Shokri, G. Theodorakopoulos, J. Y. Le Boudec, and J. P. Hubaux. Quantifying Location Privacy. In *2011 IEEE Symposium on Security and Privacy*, pages 247–262, May 2011.

[116] Reza Shokri, George Theodorakopoulos, Carmela Troncoso, Jean-Pierre Hubaux, and Jean-Yves Le Boudec. Protecting Location Privacy: Optimal Strategy Against Localization Attacks. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, CCS '12, pages 617–627, New York, NY, USA, 2012. ACM.

[117] Reza Shokri, Carmela Troncoso, Claudia Diaz, Julien Freudiger, and Jean-Pierre Hubaux. Unraveling an Old Cloak: K-anonymity for Location Privacy. In *Proceedings of the 9th Annual ACM Workshop on Privacy in the Electronic Society*, WPES '10, pages 115–118, New York, NY, USA, 2010. ACM.

[118] Herbert A. Simon. Designing organizations for an information-rich world. *International Library of Critical Writings in Economics*, 70:187–202, 1996.

[119] Yonatan Sompolinsky, Yoad Lewenberg, and Aviv Zohar. SPECTRE: A Fast and Scalable Cryptocurrency Protocol. *IACR Cryptology ePrint Archive*, 2016:1159, 2016.

[120] Yonatan Sompolinsky and Aviv Zohar. Secure High-Rate Transaction Processing in Bitcoin. In *Financial Cryptography and Data Security*, Lecture Notes in Computer Science, pages 507–527. Springer, Berlin, Heidelberg, January 2015.

[121] Yonatan Sompolinsky and Aviv Zohar. *PHANTOM*. 2018. Published: IACR Cryptology ePrint Archive, Report 2018/104.

[122] Zak Stambor. Mobile ads account for 91% of Facebook's ad revenue in Q1, April 2018. Accessed: https://www.digitalcommerce360.com/2018/04/25/mobile-ads-account-for-91-of-facebooks-ad-revenue-in-q1/.

[123] Latanya Sweeney. k-Anonymity: A Model for Protecting Privacy. *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, 10(5):557, October 2002.

[124] Cisco Systems. Cisco 2016 Midyear Cybersecurity Report, 2016. Accessed: http://www.cisco.com/c/m/en_us/offers/sc04/2016-midyear-cybersecurity-report/index.html.

[125] Douglas B Terry, Marvin M Theimer, Karin Petersen, Alan J Demers, Mike J Spreitzer, and Carl H Hauser. Managing update conflicts in Bayou, a weakly connected replicated storage system. In *Proceedings of the fifteenth ACM symposium on Operating systems principles (SOSP '95)*, volume 29. ACM, 1995.

[126] Muhammad Usman, Anteneh A Gebremariam, Usman Raza, and Fabrizio Granelli. A software-defined device-to-device communication architecture for public safety applications in 5g networks. *IEEE Access*, 3:1649–1654, 2015.

[127] Vivek Wadhwa. Laws and Ethics Cant Keep Pace with Technology. *MIT Technology Review*, April 2014.

[128] Stanley L. Warner. Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association*, 60(309):63–69, March 1965.

[129] W. Wei, X. Zhu, and Q. Li. LBSNSim: Analyzing and modeling location-based social networks. In *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, pages 1680–1688, April 2014.

[130] Marius Wernke, Pavel Skvortsov, Frank Drr, and Kurt Rothermel. A Classification of Location Privacy Attacks and Approaches. *Personal Ubiquitous Comput.*, 18(1):163–175, January 2014.

[131] Stephen B. Wicker and Kolbeinn Karlsson. Internet Advertising: Technology, Ethics, and a Serious Difference of Opinion. *Commun. ACM*, 60(10):70–79, September 2017.

[132] Ben Williams. Adblock Plus and (a little) more: Adblock Plus for iOS is finally here! Pssst, it's free!, September 2015. Accessed: https://adblockplus.org/blog/adblock-plus-for-ios-9-finally-here-and-pssst-it-s-free.

[133] James Williams. Why It's OK to Block Ads, October 2015. Accessed: http://blog.practicalethics.ox.ac.uk/2015/10/why-its-ok-to-block-ads/.

[134] Nick Wingfield and Mike Isaac. Pokmon Go Brings Augmented Reality to a Mass Audience. *The New York Times*, July 2016.

[135] Gavin Wood. Ethereum: A secure decentralized generalized transaction ledger. Technical report, 2014. Ethereum project yellow paper, EIP-150 Revision.

[136] Z. Xiao, J. Xu, and X. Meng. p-Sensitivity: A Semantic Privacy-Protection Model for Location-based Services. In *2008 Ninth International Conference on Mobile Data Management Workshops, MDMW*, pages 47–54, April 2008.

[137] Mingqiang Xue, Panos Kalnis, and Hung Keng Pung. Location Diversity: Enhanced Privacy Protection in Location Based Services. In Tanzeem Choudhury, Aaron Quigley, Thomas Strang, and Koji Suginuma, editors, *Location and Context Awareness*, number 5561 in Lecture Notes in Computer Science, pages 70–87. Springer Berlin Heidelberg, 2009.

[138] Elias Yaacoub and Osama Kubbar. Energy-efficient device-to-device communications in LTE public safety networks. In *Globecom Workshops (GC Wkshps), 2012 IEEE*, pages 391–395. IEEE, 2012.

[139] Dingqi Yang, Daqing Zhang, Longbiao Chen, and Bingqing Qu. NationTelescope: Monitoring and visualizing large-scale collective behavior in LBSNs. *Journal of Network and Computer Applications*, 55:170–180, September 2015.

[140] Dingqi Yang, Daqing Zhang, and Bingqing Qu. Participatory Cultural Mapping Based on Collective Behavior Data in Location-Based Social Networks. *ACM Trans. Intell. Syst. Technol.*, 7(3):30:1–30:23, January 2016.

[141] Hui Zang and Jean Bolot. Anonymization of Location Data Does Not Work: A Large-scale Measurement Study. In *Proceedings of the 17th Annual International Conference on Mobile Computing and Networking*, MobiCom '11, pages 145–156, New York, NY, USA, 2011. ACM.

[142] Keshu Zhang, Haifeng Li, Kari Torkkola, and Mike Gardner. Adaptive Learning of Semantic Locations and Routes. In *Proceedings of the 1st International Conference on Autonomic Computing and Communication Systems*, Autonomics '07, pages 3:1–3:10, ICST, Brussels, Belgium, Belgium, 2007. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

[143] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. Mining Interesting Locations and Travel Sequences from GPS Trajectories. In *Proceedings of the 18th International Conference on World Wide Web*, WWW '09, pages 791–800, New York, NY, USA, 2009. ACM.

[144] Kathryn Zickuhr. Location-Based Services. Technical report, September 2013. [Online]. http://www.pewinternet.org/2013/09/12/location-based-services/.