# An Investigation into Location Privacy of Navigation Applications: A Systematic Literature Review

Shane Ahern

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

M.Sc. in Management of Information Systems

2019

## Declaration

I declare that the work described in this dissertation is, except where otherwise stated, entirely my own work, and has not been submitted as an exercise for a degree at this or any other university. I further declare that this research has been carried out in full compliance with the ethical research requirements of the School of Computer Science and Statistics.

Signed: _____

## Permission to lend and/or copy

I agree that the School of Computer Science and Statistics, Trinity College may lend or copy this dissertation upon request.

Signed: _____

## Acknowledgements

I would like to thank my supervisor Melanie Bouroche for her help, advice and guidance during the writing of this dissertation.

I would also like to thank my family, friends and colleagues for their patience and support during this endeavour.

## Abstract

The rapid technological advancements in location tracking technologies and devices have revolutionized how people consume navigation applications. Historically, navigation was achieved through the use of physical maps and compasses. The emergence of internet connectivity in the 1990s led to map digitization which was the precursor of modern-day navigation applications. Today, navigation applications have become integrated into people's daily lives through the use of smartphones. The smartphone has transformed navigation making travel planning frictionless. The technology used is capable of tracking the user location to an extremely high precision which enables a superior standard of navigation accuracy. While this is advantageous to consumers, it also raises major concerns around the area of consumer privacy, specifically location privacy. Many navigation applications are provided by third-party service operators who capture and continually track user location. This paper uses a Systematic Literature Review to examine how navigation applications calculate travel routes and track user locations. The study analyses the location tracking technologies, devices and underlying systems used. The focus of the paper is to examine the impact navigation applications have on user privacy. The paper identifies the threats to consumer privacy and examines the effectiveness of the privacy protections in place.

# Table of Contents


1 Introduction ..... 1
1.1 Navigation Applications ..... 1
1.2 Privacy ..... 4
1.3 Research Objectives ..... 6
1.4 Scope ..... 7
1.5 Benefits of the Research ..... 7
1.6 Dissertation Structure ..... 8
2 Literature Review ..... 9
2.1 Location Technologies ..... 9
2.2 Navigation Devices ..... 18
2.3 Multi Modal Transportation Systems ..... 22
2.4 Privacy Threats ..... 27
2.5 Privacy Protections ..... 34
3 Research Methodology ..... 45
3.1 Research Design ..... 45
3.2 Systematic Literature Review ..... 45
3.3 Purpose of Literature Review ..... 46
3.4 Research Question and Protocol ..... 47
3.5 Literature Search Process ..... 47
3.6 Practical Screening for inclusion ..... 49
3.7 Quality Appraisal for Exclusion ..... 51
3.8 Data Extraction ..... 51
3.9 Synthesize Studies ..... 52
3.10 Writing the review ..... 52
3.11 Limitations of Methodology ..... 53
3.12 Lessons Learnt ..... 53
4 Conclusion ..... 54
4.1 Answering the Research Questions ..... 54
4.2 How Navigation Applications Work ..... 54
4.3 Privacy threats ..... 57
4.4 Privacy Protections ..... 58
4.5 Limitations of Research ..... 60
4.6 Future Research ..... 60
References ..... 61
5 Appendices ..... 67
5.1 Appendix 1: Navigation Application Comparison (Manalo, 2019) ..... 67
5.2 Appendix 2: Smartphone 3rd party libraries (Lin, Liu, Sadeh & Hong, 2014) ..... 71
5.3 Appendix 3: Systematic Literature Review (Okoli, 2015) ..... 72

## List of Figures and Tables

Figure 1: Most Popular Navigation Applications (Panko, 2018) 2
Figure 2: Reason for choosing Navigation Application(Panko, 2018) 3
Figure 3: Google Timelines Product (Essl, 2017) 5
Figure 4: Google Timelines Daily View (Novet, 2015) 6
Figure 5: GPS Trilateration (GISGeography, 2019) 10
Figure 6: GPS Horizonal Position Error Histogram (Hughes, 2017) 12
Figure 7: GPS Vertical Position Error Histogram (Hughes, 2017) 12
Figure 8: GPS Signal Spoofer Threat (Jafarnia-Jahromi et al., 2012) 13
Figure 9: Radiation Pattern Illustration (Bates, 2015) 14
Figure 10: Hybrid Positioning (P. Nath et al., 2015) 17
Figure 11: Positioning Types (L. Chen et al., 2017) 18
Figure 12: Smartphone Users Globally (Kooistra, 2018) 20
Figure 13: Post public transport switch(Abou-Zeid et al., 2012) 23
Figure 14: Super Travel API design (Evangelatos et al., 2017) 24
Figure 15: Valuing Person Data Framework (World Economic Forum, 2017) 29
Figure 16: Threat Model (Freudiger et al., 2011) 31
Figure 17: Mobile Data Traffic Forecast 2016-2021 (Cisco, 2016) 32
Figure 18: Attacker Knowledge Attacks (Wernke et al., 2014) 34
Figure 19: Investment for achieving data protection compliance (Tankard, 2016) 41
Figure 20: Privacy Nudge Screenshot (Almuhimedi et al., 2015) 42
Figure 21: Position Dummies Concept (Kasori & Sato, 2015) 43
Figure 22: Cloaking Region (Kasori & Sato, 2015) 44
Figure 23: Paper Selection Flow 50
Figure 24: Systematic Mapping Process (Petersen et al., 2008) 50

Table 1: Table of search term and number of hits 49
Table 2: Location Tracking Methods Summary 55
Table 3: Navigation Device Summary 56

## List of Abbreviations

| | |
|---|---|
| API | Application Programming Interface |
| GDPR | General Data Protection Regulation |
| GIS | Geographic Information System |
| GNSS | Global Navigation Satellite System |
| GPS | Global Positioning System |
| HAR | Human Activity Recognition |
| IIN | Internet-based indoor Navigation |
| LBS | Location Based Services |
| LPPM | Location Privacy Preserving Mechanisms |
| PND | Portable Navigation Device |
| SLR | Systematic Literature Review |
| TTFF | Time to First Fix |
| WPS | Wi-Fi Positioning System |

# 1 Introduction

This chapter introduces the two main topics of the study - Navigation applications and Privacy. Additionally, the chapter outlines the research objectives, scope, benefits of the study and concludes by outlining the roadmap of the paper.

## 1.1 Navigation Applications

Navigation is a term which is defined by the Oxford Dictionary (2019b) as the “the process or activity of accurately ascertaining one's position and planning and following a route”. Humans have used navigation throughout history for movement and travel planning. Maps are the foundation to travel planning; they depict the landscape of a region representing the objects and landmarks of the area allowing the map observer to construct their optimal travel route along the mapped roads and rivers. Historically maps were concrete documents which allowed readers to physically plot and draw their route. The rapid evolution of technology and the global emergence of internet connectivity in the 1990s led to the beginning of map digitization. MapQuest launched the first ever internet-based mapped service in 1996 (Harlan, 2015) which allowed any user with an internet connection to create digital maps online which could be distributed across the world. As the internet became more prominent an increasing number of physical maps were digitized allowing them to be viewed and digested on technological devices such as computers and mobile phones.

The internet has completely transformed travel planning from the previous time-consuming task of plotting and mapping out the best route on a physical map to today’s automated solutions which make travel planning frictionless. Today, most travel planning activities are consumed on internet-based solutions. Panko (2018) in a recent survey of 511 participants found that 77% of them regularly utilise navigation applications through their mobile phones. Internet-based mapping solutions allow users to plan the optimal route to their destination by inputting their current location and desired destination location into the application which will then automatically generate travel routes which meet the user’s specification. Ambrose, Bukovsky, Sedlak, and Goeden (2009) discussed the complexity faced by route planners to create the optimal route for each user. The optimal route could mean different things for different people depending on their motivation. To cater for different consumer factors, navigation applications recommend several routes to the user from which they can choose their preferred route. The routes take into consideration information about traffic congestion, form of transportation and any potential costs which may be incurred along each route.

Navigation applications can be spilt into two specific application types; travel planners and travel assistants. The distinction is important from a privacy stand point due to the significant difference between the real time location data tracked by each application. Travel planners are applications which plot the best route currently available based on two inputs from the user, their starting location and their desired destination location. The data captured and tracked is limited to the inputs provided by the user; they are not aware of and do not track the current location of the user.

Travel assistants leverage the significant use of smart mobile devices and the ever-expanding mobile network connectivity to provide a real time application that tracks the current location of the user and prompts new routes that would get the user to their destination more quickly based on insights which have been gained since the journey began. These insights can be gained from analysing current route conditions which include traffic congestion and potential route delays due to road closures and traffic accidents.

There are several advantages to the end user of using a travel assistant which provides real time information and re-routing capabilities. However, the user location is continually tracked throughout the journey which leads to location privacy concerns which are discussed in detail in chapter 2.4.

### *1.1.1 Navigation Applications Market*

Panko (2018) market survey on navigation applications in 2018 found that Google Maps was by far the leading application in the Navigation App industry with a 67% share of the market followed by Waze with 12%, Apple Maps with 11% and MapQuest with 8%.

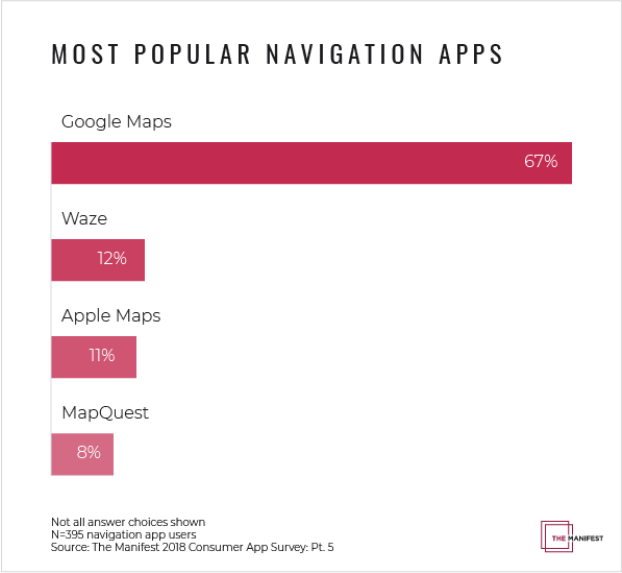


*Figure 1: Most Popular Navigation Applications (Panko, 2018)*

The survey also asked users their reason for choosing their favourite navigation application with a leading 25% claiming clear directions as their reason while 20% choose preferred features and user-friendly design/interface with 17% choosing best directions for non-drivers and 14% stating they have never used another navigation app.

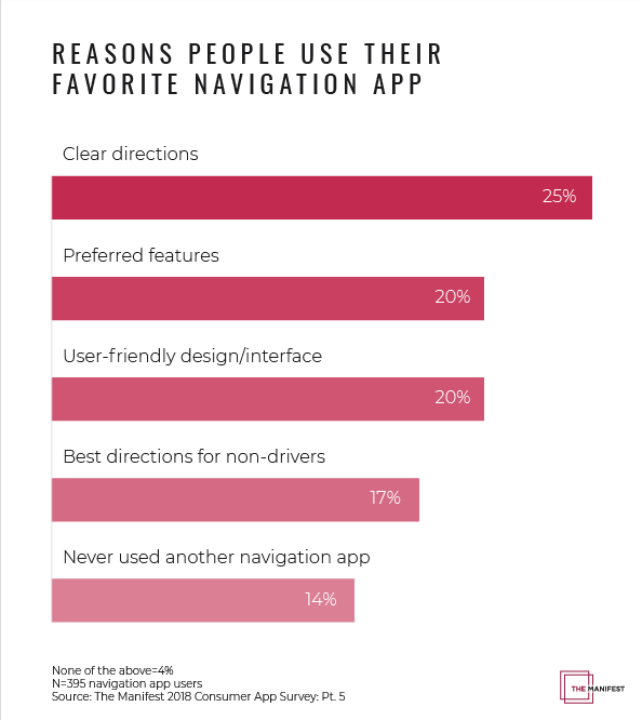


*Figure 2: Reason for choosing Navigation Application(Panko, 2018)*

Gadget Hacks an electronics consumer website compared the functionality of the four leading navigation applications; Google Maps, Waze, Apple Maps and MapQuest across 41 different criteria to determine the most feature rich application. The comparison found that Google Maps had a significant advantage over the competition with comprehensive functionality across the varied criteria. The full results can be seen in Appendix 1 (Manalo, 2019). With such dominance in terms of functionality it is little surprise that google maps is the most widely used navigation application.

### *1.1.2 Navigation Applications, why are they needed?*

*"Transportation is the center of the world! It is the glue of our daily lives. When it goes well, we don't see it. When it goes wrong, it negatively colors our day, makes us feel angry and impotent, curtails our possibilities"*

Robin Chase, CEO of Zipcar

Travel is ever present in people's daily lives; it pertains to everything that is associated with the movement of people. People travel for a wide variety of reasons which include commuting to work or travelling for leisure or pleasure. Transportation can sometimes be problematic with delays caused by traffic congestions or issues with transportation leading to increased stress levels for the people involved (Crotts & Zehrer, 2012). Navigation applications can help ease the stressful burden of travelling by providing real time information which enables the user to plan and track their destination while adapting to ongoing route conditions. Schmitt, Currie, and Delbosc (2014) carried out research on Melbourne's public transportation system to measure the satisfaction score and overall travel experience of first-time commuters against experienced commuters who had previously taken the route. The study found that unfamiliar travellers rated the experience of their journey as far more negative than seasoned familiar travellers. The study found that the two key attributes pertaining to the negativity of unfamiliar travel were navigation and emotional state during the trip. Navigation in this context refers to the traveller's experience while on public transportation; attempting to figure out their current location in the context of their trip. Increased anxiety was observed in unfamiliar travellers which may be as a result of travelling in an unknown area not knowing the stop at which they are meant to disembark. The study highlights the importance of accurate navigation systems which would ease the anxiety felt by unfamiliar travellers enabling them to track their location in real time.

## 1.2 Privacy

Privacy is defined in the oxford dictionary as "A state in which one is not observed or disturbed by other people" (Oxford Dictionary, 2019b). This definition conveys the physical attributes to privacy; that people have the right to not be interfered with by other people. However, in the information technology era privacy has evolved. People are no longer only observed or disturbed by other people; technological information systems are capable of amassing huge amounts of data on users which is captured through sensor-rich devices which have become prevalent across the world. A new term 'Information Privacy' has emerged with the rise of information services. Information privacy is the term used to describe the relationship between information system data collection and processing policies and the consumer's expectation of privacy. Westin (1967) was one of the initial researchers to discuss information privacy, describing it as the 'claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others'. Today, it is difficult to understand the extent to which personal information is captured, processed and used by third-party service providers. For many services, consumers are required to accept their

privacy policy before using the service. The privacy policy is a legal document which outlines the data management processes of the organisation; how the service captures, processes and discloses the user information. Navigation applications require location data on the user's precise location and destination location to provide a mapping service which will generate a route for the consumer. As such navigation applications process a significant amount of sensitive location information related to the user. The sheer volume of data captured is highlighted by Google's maps 'Timelines' product which is a visual representation of the user's timeline based on location data captured and processed through the use of location enabled devices which are linked to the user's google account.

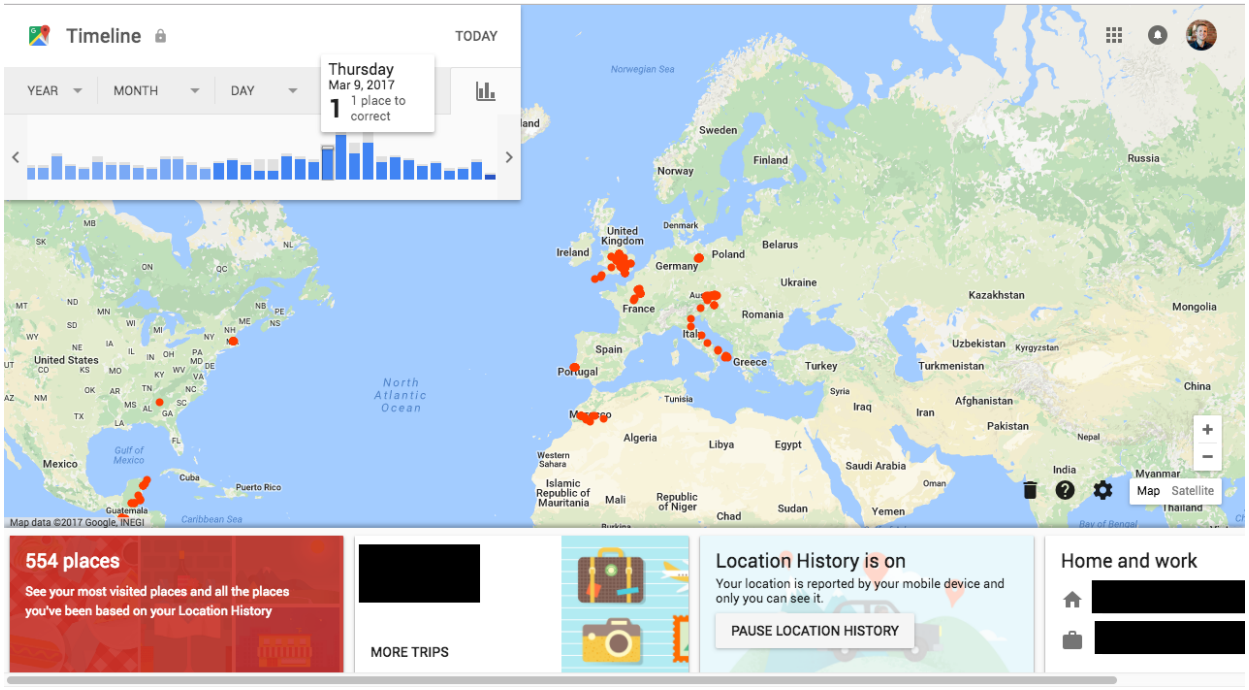


*Figure 3: Google Timelines Product (Essl, 2017)*

Figure 3 illustrates Google's Timelines product. The default view of the product provides a global overview of the places the user has been which are dotted across the map of earth. In the bottom left corner, the tool displays the number of places that have been identified based on location data provided. Clicking into this section provides an overview of the places, number of days visited and the last date the user was there. The places are ranked by the number of visits. On the top left corner, the tool provides a date picker which allows the user to drill into detailed location information from a selected date. The information provides a detailed overview and visual map representation of the user's day, including the places they have visited, the time they were there and how long it took them to get to each destination. The detailed day view is displayed in figure 4 below.

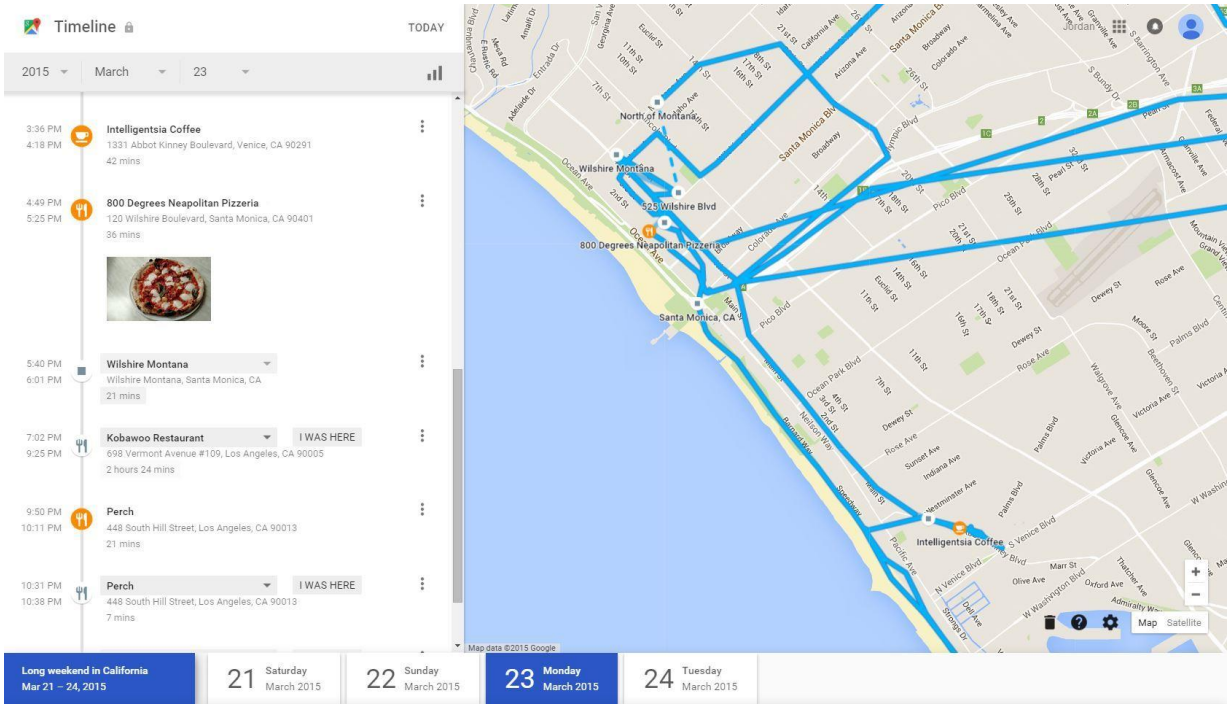


*Figure 4: Google Timelines Daily View (Novet, 2015)*

Worryingly, the timelines product is switched on by default for all google accounts (Fowler, 2018). The volume of location data being captured, processed and presented by google is staggering which raises location privacy concerns for consumers. Clarke and Wigan (2011) in their research titled '*You are where you've been*' identified location tracking technologies as a direct threat to users' privacy. The researchers indicated that the volume of data stored on each user about their present and past locations allow the observers to impute aspects of the person's behaviour and habits raising the risk of user profiling.

## 1.3 Research Objectives

The aim of this study is to analyse the impact navigation applications have on consumer's privacy. The study will investigate how navigation applications work to understand the potential y threats to privacy raised by using such applications. The study will also discuss the privacy protections in place which are used to defend consumers against such threats to their privacy.

The study aims to:

1. Gain an understanding of how navigation applications work.
2. Identify the potential privacy threats imposed through the use of navigation applications.
3. Identify the privacy protections in place to protect users of navigation applications

## 1.4 Scope

The study used a systematic literature review approach to critically analyse privacy in navigation applications. No primary data was collected in this study, the study used secondary data from a wide range of literature to provide an in-depth insight into the topic. The reasons for using this approach will be discussed in detail in Chapter 3, Research Methodology.

## 1.5 Benefits of the Research

As previously discussed, the use of navigation services is ever increasing. The rise of ubiquitous computing has led to increases in the volume of personal data captured and processed by third-party organisations. The aim of this paper is to uncover the privacy concerns that are exposed through the sharing of location information with navigation service providers. The study highlights these threats to privacy and discusses the technical and constitutional protections in place to defend users against such imposed threats.

The aim of the study is to expand on the existing knowledge base of privacy in the context of navigation applications. The study may be considered as a resource for privacy regulators and policy makers who wish to gain a better understanding in the navigation applications field. Additionally, the study may be useful for consumers of navigation applications to gain a better understanding of the processes involved in generating their navigation routes and to understand how navigation service providers process their sensitive location data.

## 1.6 Dissertation Structure

The purpose of this section is to provide the reader with the roadmap for how the dissertation is laid out including a description of the key points contained within each chapter.

*Chapter 1: Introduction*

This chapter provides the reader with an overview and background information on the context of the study. The aim and scope of the research is defined.

*Chapter 2. Literature Review*

This chapter examines the literature on location privacy in the navigation applications market. The chapter begins by analysing how navigations systems work and what devices are used to consume these services. Furthermore, the privacy threats of location data and privacy defences are discussed in detail.

*Chapter 3. Research Methodology*

This chapter outlines the methodology used while undertaking this study. The chapter discusses the strengths and weaknesses of the approach and the reasons for choosing the approach.

*Chapter 4. Conclusion*

The study concludes by emphasising the key findings of the research and how it adds to the body of knowledge on location privacy. It discusses the limitations of the research and proposes potential further research on the topic.

# 2 Literature Review

This chapter will assess and analyse the current knowledge of the topic through a systematic review of literature. The chapter begins by discussing the enabling innovation behind location tracking: Location tracking technologies and navigation devices. The chapter will then continue by discussing multi-modal transportation systems to gain an understanding of the inner workings of navigation applications. Finally, the chapter will conclude by discussing privacy threats and protections which are in place for consumers of navigation applications. The chapter is therefore divided into five parts:

1. Tracking Technologies
2. Navigation Devices
3. Multi-Modal Transportation Systems
4. Privacy Threats
5. Privacy Protections

## 2.1 Location Technologies

This section will discuss geographical information systems which store and manage spatial data which is used by all other location tracking technologies. The core technical details of four different location tracking technologies: *Global Navigation Satellite System, Mobile Network Location Tracking, Wi-Fi Positioning System and Hybrid Positioning System* will be investigated and the advantages and disadvantages of using each approach will be outlined.

### *2.1.1 Geographical Information System*

Geographical Information System commonly abbreviated to GIS is the framework for managing all spatial data, data that is related to the occupation of space. The term geographical in the context of information systems implies that the data objects stored in the system are known and can be calculated in terms of longitude and latitude in a topographical format (Heywood, Cornelius, & Carver, 2011). GIS systems are used to display and manipulate digital maps and images related to a geographical position. Geographical information system is the foundation for providing navigation services, which utilise the geographical position data to generate the route information to get the end user to their destination through mapped roads and rivers. The most widely used web-based GIS as discussed in chapter 1.1.1 is google maps which integrates with Google Earth to represent the spatial information for planet earth. GIS systems differ from the other

systems outlined in this section as their purpose is to store spatial data of the earth. Each of the other location tracking systems utilise the data stored in the Geographical Information System to identify the device's position on earth.

### *2.1.2 Global Navigation Satellite System*

Global Navigation Satellite System commonly abbreviated to GNSS is a satellite-based navigation system. The U.S Department of Defence created the first and most well-known GNSS system; Global Positioning Systems (GPS) in 1973 "to establish, develop, test, acquire, and deploy a spaceborne positioning system" (Hofmann-Wellenhof, Lichtenegger, & Collins, 2012). As of 2019 there are three GNSS operational:

1. Global Positioning System (GPS), United States
2. Global'naya Navigatsionnaya Sputnikovaya Sistema (GLONASS), Russia
3. Galileo, European Union

Global Navigation Satellite Systems (GNSS) use a minimum of 24 operational GPS satellites that orbit earth at an altitude of 20,200 kilometres and emit a GPS radio frequency. The satellites are arranged in a constellation which work together to ensure location availability and accuracy (United States Government, 2019). GNSS satellites use an extremely stable atomic clock which is synchronized across all satellites to maintain the true time in the constellation. The position of each GPS satellite is always known to a very high accuracy. Devices which are capable of interpreting GPS signals known as receivers can calculate their location through a process called trilateration. The receiver device will receive multiple radio frequencies from multiple satellites, using the time It took to retrieve the message and the exact location of the satellites at the time of transmission the device location can be determined.

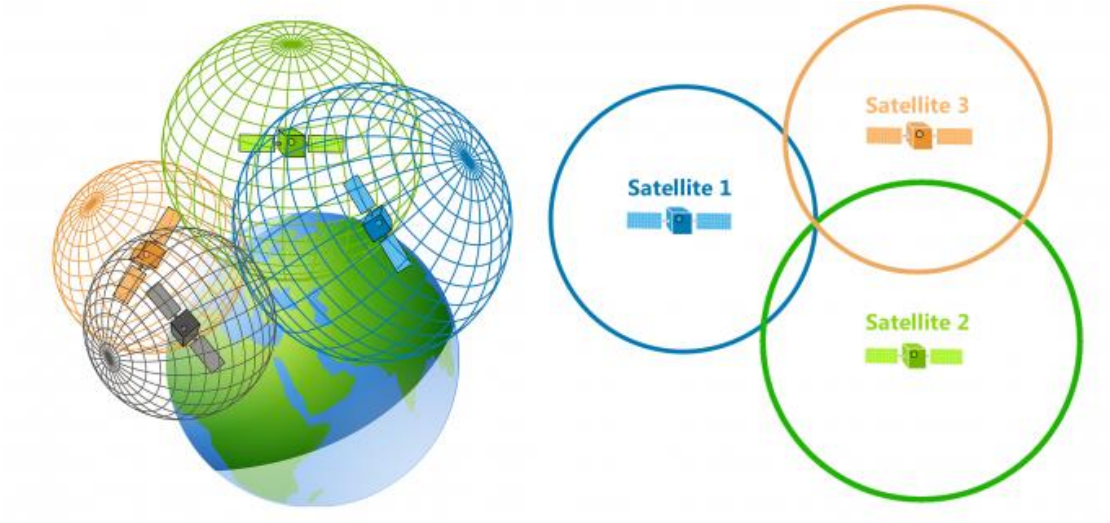


*Figure 5: GPS Trilateration* (GISGeography, 2019)

Figure 5 above depicts trilateration, the process of determining location based on the distance of the receiver from multiple satellites. The intersection of the three circles illustrated above gives the location of the receiver; The longitude, latitude and altitude will be calculated. In order to provide a location, the receiver needs to know the distance between the receiver and each of three or more satellites and the exact time based on the atomic clock of each satellite. Less than three satellites would make it impossible to get an accurate location as the intersection of two radio frequencies would result in the radii intersecting at two different points making it impossible to determine which is the location of the receiver.

Time to First Fix (TTFF) is a measure of the performance of location tracking, it is a calculation of the time taken by a receiver to calculate the initial position of the device. There are three main scenarios which impact the performance and time taken to receive the location information of the device: cold, warm and hot. Cold start refers to the location calculation of the object which does not have any location data stored on the receiver. It must complete a full sky scan to systematically query all possible satellites. As such cold location queries are the slowest to perform. Warm start indicates that the receiver has relatively accurate time and position data stored, requiring the device to only retrieve detailed orbital information from the recently queried satellite. Hot starts are the fastest location lookup, the receiver contains all the necessary data from a warm start are stored with the addition of accurate location and time data. (Paonni, Anghileri, Wallner, Avila-Rodriguez, & Eissfeller, 2010)

The Federal Aviation administration of United States performance report (Hughes, 2017) on their Global Positioning System found that it was highly performant with an accuracy of within 1.891 metres of horizontally and 3.872 metres of vertically 95% of the time. The vertical and horizontal error probability graphs are illustrated below.

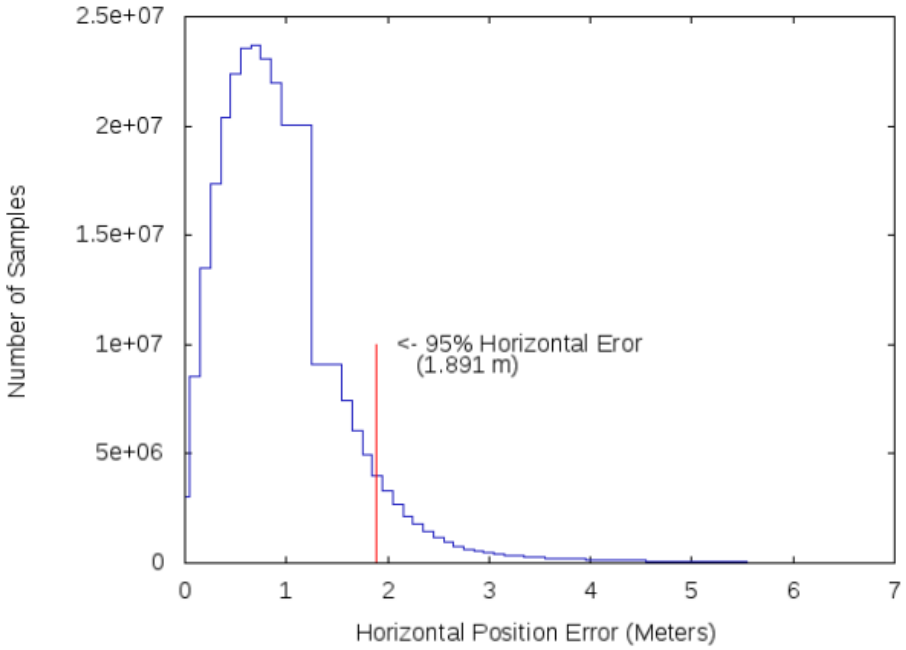


*Figure 6: GPS Horizonal Position Error Histogram (Hughes, 2017)*

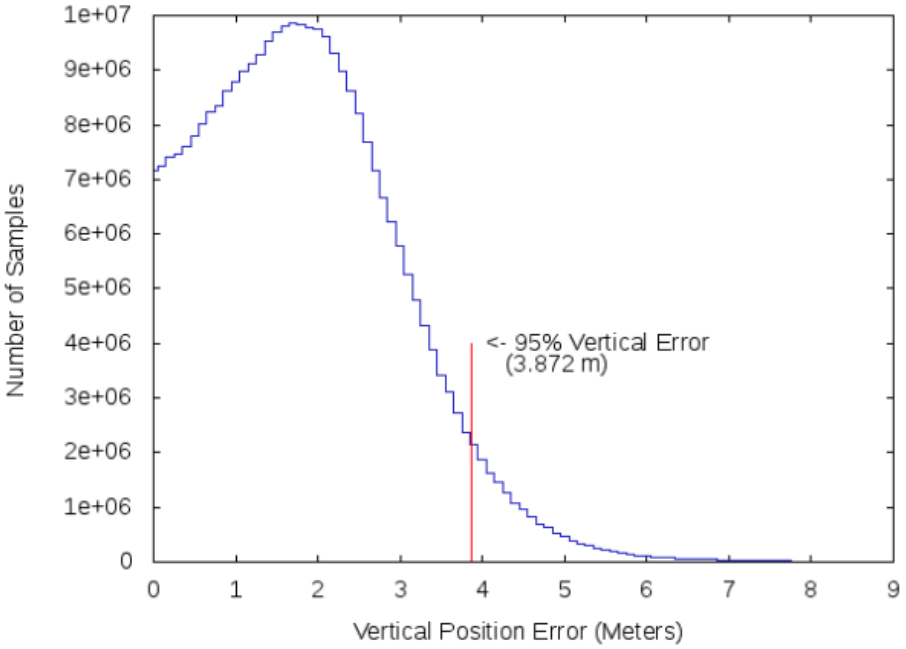


*Figure 7: GPS Vertical Position Error Histogram (Hughes, 2017)*

Per Enge (2017) in the Stanford podcast 'How safe and secure is GPS' identified the two key threats to GPS as Interference and Spoofing. Interference is the overwhelming of the receiver commonly referred to as signal jamming which prevents the signals from reaching the retriever. Spoofing is the imitation of signals; sending out invalid and fake GPS signals to mislead consumer devices. Jafarnia-Jahromi, Broumandan, Nielsen, and Lachapelle (2012) further discussed these threats to GNSS highlighting in-band interference vulnerabilities that are easy to target due to the low power signal emitted by

the satellites. Figure 8 below portrays the threat of signal imitation in which the Spoofer device sends an altered signal to the target receiver, the researchers described this attack as ‘potentially significantly more menacing than jamming since the target receiver is not aware of the threat’.

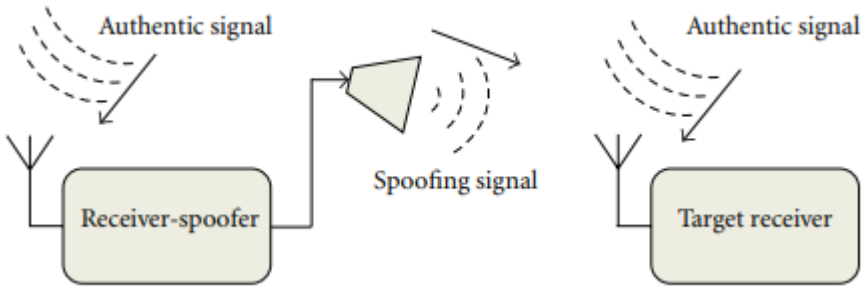


*Figure 8: GPS Signal Spoofer Threat (Jafarnia-Jahromi et al., 2012)*

An argument could be made that we are over reliant on GNSS systems for navigation. There are an extremely limited number of GNSS systems operational which are operated and governed by two countries - The United States and Russia and one consortium of countries the European Union. Due to the monumental costs associated with launching a constellation of satellites there is little to or no competition in the industry with the quality of service and standards directly linked to Government initiatives and funding. GPS, the initial GNSS system, was initially designed for military operations and logistics and has now become a ubiquitous utility to our daily lives which we depend on for navigation. Additionally, the GPS has become the core technology guiding our transportation industries including aviation and shipping which would cease to function effectively if GNSS Systems were to break or if their accuracy diminished.

The flow of information for GPS to receivers is unidirectional, there is no communication from the receiver device that the GPS radio frequency had been received. As such, GPS has built- in privacy as the user’s location is calculated on the device, no location data is sent back to the transmitting satellites. (Beresford & Stajano, 2003)

### *2.1.3 Mobile Network Location Tracking*

Mobile network providers use ground-based antennas and radio towers to create a wireless communications network that can provide mobile telecommunication capabilities to a region. Network based location tracking for smart phones can determine the location of the device by leveraging the service provider's network. Mobile network location tracking can be spilt into three types:

1. Cell Identification
2. Cell Tower Triangulation
3. Multilateration Localization

Cell identification is the lowest accuracy location tracking type, it uses the signal strength patterns of mobile network towers to identify the location of the device. The technique known as radiation pattern works off the precedence that the Antennas emit frequencies in a known pattern and strength, thus a location estimate for the receiver can be made based on the strength of the mobile signal received from an identifiable cell tower whose location is known and stationary. Figure 9 below illustrates the concept of radiation patterns which extend in a known pattern from the source antenna.

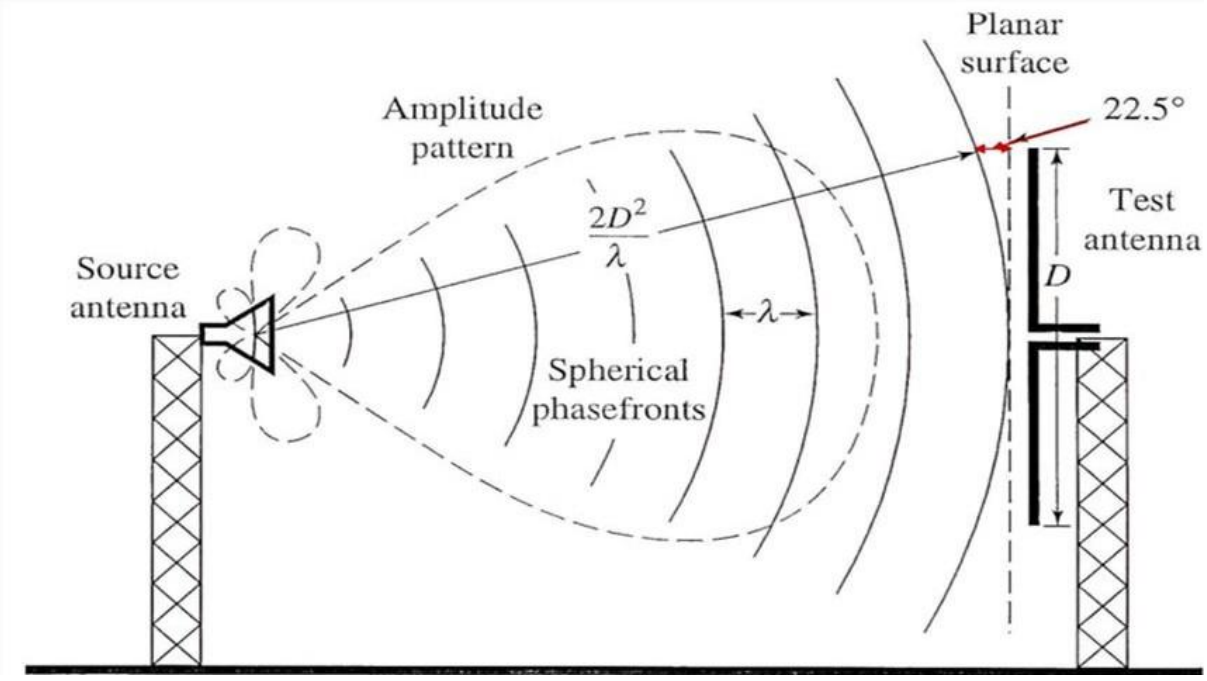


*Figure 9: Radiation Pattern Illustration (Bates, 2015)*

Cell tower triangulation like Global Navigation Satellite Systems uses triangulation to determine the position of the receiver (Yang, Varshavsky, Liu, Chen, & Gruteser, 2010). The distance between the cell towers and the receiver is calculated based on the time taken to receive the signal. This calculation is done for each cell tower in the area for which the device is receiving transmission. Using the distance between known cell tower

locations as a reference the triangulation technique which is explained in detail in the GNSS section enables the receiver location to be identified. The accuracy is Cell tower triangulation depends on the number of cell towers in range of the device. As such in urban areas the accuracy would be much improved with large numbers of radio towers within range. In rural areas the number of radio towers would be significantly less which would greatly reduce the location accuracy received.

(Zhou, Jun, & Lamont, 2012) identified Multilateration localization as a key technique used to identify the location of nodes on a wireless sensor network. The technique is based on distance measurements between an anchor node and the receiver on a network. Anchor nodes are simply sensors whose location is stationary and known on the wireless network. You could argue that cell tower triangulation and multilateration localization work off the same principle as GPS. They all use distance between known locations and the receiver to calculate its location. Network based location tracking techniques preceded GPS tracking on mobile devices as GPS receivers were initially customised and limited devices utilised primarily by the US military.

Similar to GPS, Mobile Network location tracking is unidirectional, the location of the device is not relayed back to the antennas making location privacy an innate feature.

### *2.1.4 Wi-Fi Positioning System*

As transportation navigation systems have become more ubiquitous the research community have identified the need for indoor navigation systems to help users find their way through complex indoor structures and areas. Wi-fi positioning system (WPS) which use information from nearby wi-fi access points to locate the receiver are commonly used for in door navigation systems. Wi-fi is the preferred indoor positioning system as GPS signals may be weak due to being obstruction by the environment. Additionally, Wi-fi based location systems have an improved time to first fix for indoor systems as the distance between the wi-fi access points and the receiver is much shorter than satellite systems.

(Han, Jung, Lee, & Yoon, 2014) created an indoor wi-fi based system in the COEX Seoul in 2010. The researchers outlined their seven-step process in building the wi-fi based indoor navigation system:

1. Analyse the Wi-Fi access points available in the area
2. Design Goals Set up
3. Indoor Map Drawing
4. Wi-fi Radio Map Construction

5. System Build-Up
6. System Testing
7. Service Launching and User Feedback

Two of the key steps in the creation of the system were steps three and four; the indoor map drawing and wi-fi radio map construction. These steps entailed the creation of the internals map of the building along with Wi-fi fingerprint data collection across the entire building. Once completed the researched had created the Wi-Fi radio map which they then installed a localization and navigation engine on top.

(Zeinalipour-Yazti, Laoudias, Georgiou, & Chatzimilioudis, 2018) cited location privacy concerns with Internet-based indoor navigation (IIN) systems which know the user location throughout their time connected to the Wi-Fi system. The researchers categorised privacy as all calculations are done on the carrier device which has its own built in localization algorithm which uses sensors locally. Network-based processing offers no privacy to the consumer; the approach requires use of the IIN service where contains the localization algorithm. As such the position of the device is always known on the network by the IIN service provider.

### *2.1.5 Hybrid Positioning System*

Each location tracking technology has limitations when used individually to locate devices:

- GPS signals may be obstructed by the environment
- Mobile network localization is heavily dependent on number of radio towers in a specific region
- Wi-fi signal range is extremely limited.

Wan, Wan Bejuri, Mohamad, Mohd, and Sapri (2011) identified the need to use multiple location tracking types to determine the location of smart phones which overcome the limitations of each type. Hybrid positioning, the combination of different location tracking technologies into an integrated solution, improves the accuracy of determining the location of mobile devices across all environments. (Ratsameethammawong & Kasemsan, 2010) discussed the use of a combination of GPS, Wi-Fi and Mobile network technologies to determine the location of mobile phones. The researchers found that using a combination of the three technologies enabled a more accurate location to be determined. Similarly, (P. Nath, Parija, Sahu, & Singh, 2015) found that time to first fix (TTFF) for hybrid positioning systems was drastically improved using hybrid positioning systems. Cell towers have GNSS receivers built in which constantly pulls satellite information. The nearby cell tower to the mobile device can provide this GNSS information enabling a

'warm' TTFF' lookup as some of the required GNSS data is known. Hybrid positioning systems are the standard system used in mobile devices as it offers a wide range of capabilities to locate devices in many different terrains. Figure 10 below illustrates hybrid positioning using nearby cell tower data to reduce TTFF.

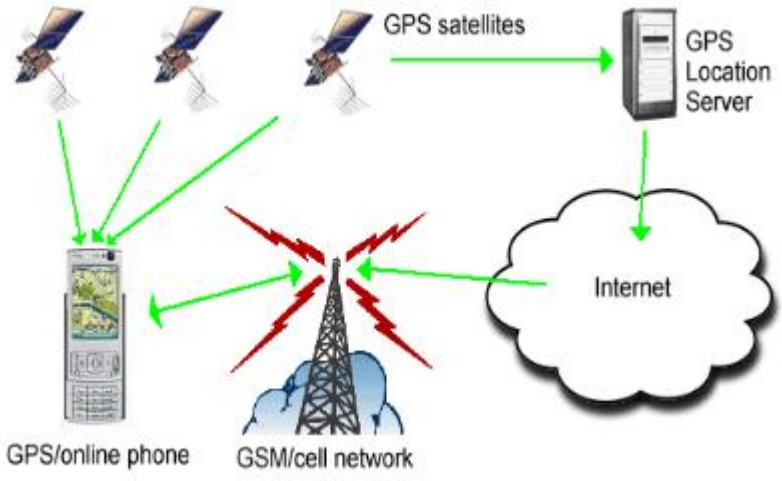


*Figure 10: Hybrid Positioning (P. Nath et al., 2015)*

### *2.1.6 Location Tracking Methods Summary*

L. Chen et al. (2017) classified the location positioning technologies into three main classes:

1. GNSS Positioning
2. Assisted-GNSS and Cloud-GNSS Positioning
3. Non-GNSS Positioning

The research evaluated the security and privacy risks of using each tracking type in the context of internet of things devices; they found that GNSS positioning alone was the most private positioning technology as the user location is calculated on the device itself. The study also found that Hybrid positioning systems and Non-GNSS positioning systems such as Wi-fi raised privacy risks as the user location data was shared with external parties. Figure 11 illustrates this point, Non-GNSS and Assisted GNSS utilise a location aggregator and service provider to determine the position while GNSS solely uses the device.

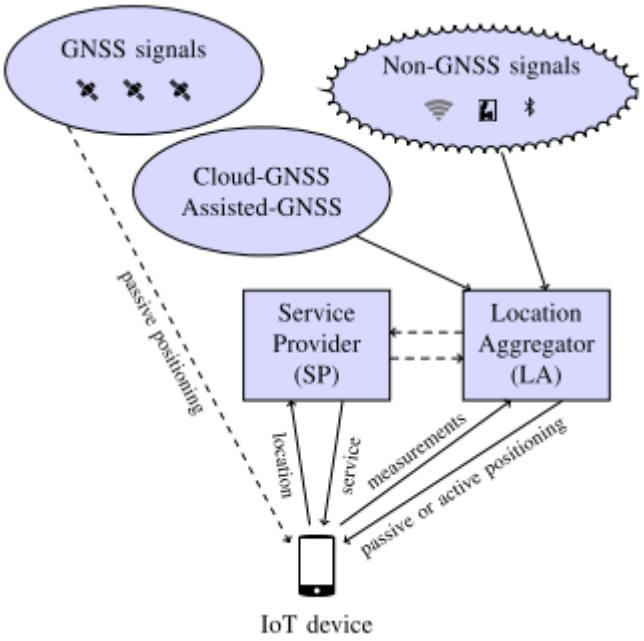


*Figure 11: Positioning Types (L. Chen et al., 2017)*

## 2.2 Navigation Devices

Navigation devices are simply devices which can retrieve location data from Global Navigation Satellite System (GNSS). This section will outline the three primarily used consumer devices utilised for navigation:

1. Portable Navigation Devices
2. In-Dash Navigation
3. Smartphones

### *2.2.1 Portable Navigation Devices*

Portable Navigation Devices (PND) commonly referred to as Sat Navs were the initial consumer facing navigation system which has two core capabilities: Location Positioning and Navigation. PND have GPS receiver capabilities which allow the device to use GPS location tracking to determine its position. PND also have built in navigation algorithms which generate the route to the user's destination. Portable Navigation devices' sole purpose is navigation, which has led to their demise in the industry. The rise of GPS enabled smart phones has rendered Portable Navigation Devices obsolete with a significant decline in their share of the navigation device market.

### *2.2.2 In-Dash Navigation*

In Dash navigation systems are navigation systems which are built into the vehicle itself. The navigation system is displayed as part of the vehicle dashboard presented to the driver. In-dash navigation systems use a similar principle as Portable Navigation Devices; it tracks the vehicle position using GPS and calculates the user's route based on an in-built algorithm. According to market research carried out the In-Dash navigation market is expected to reach revenues of cost $20 billion by the end of 2022 (PersistenceMarketResearch, 2017).

It could be argued that Portable and In-dash Navigation Devices protect the privacy of the consumer. They use their own built in algorithm to generate the required route ensuring that the consumers data is confined within the device itself. This contrasts with the use of third-party services such as Google maps which is commonly used to generate the route for the user in other navigation applications.

### *2.2.3 Smartphones*

Smartphones are small portable devices which have significant telecommunications and computational capabilities. The smartphone combines the functionality of many devices into one single piece of hardware which is pocket sized and easily carried. The hyper convergence of disparate devices into one powerful handheld device has led to the Smartphone becoming one of the most widely used devices. Many consider smartphones to be critical to everyday life enabling people to stay connected to the internet and each other from anywhere in the world.

Recent market research carried out by Newzoo (Kooistra, 2018) found that in 2018 there were 3 billion active smartphone users globally which is predicted to rise to 3.8 billion users by 2020. To put those colossal figures into context, just shy of 40% of the world population had smartphones in 2018 and this is expected rise to over 50% of the global population by 2020.

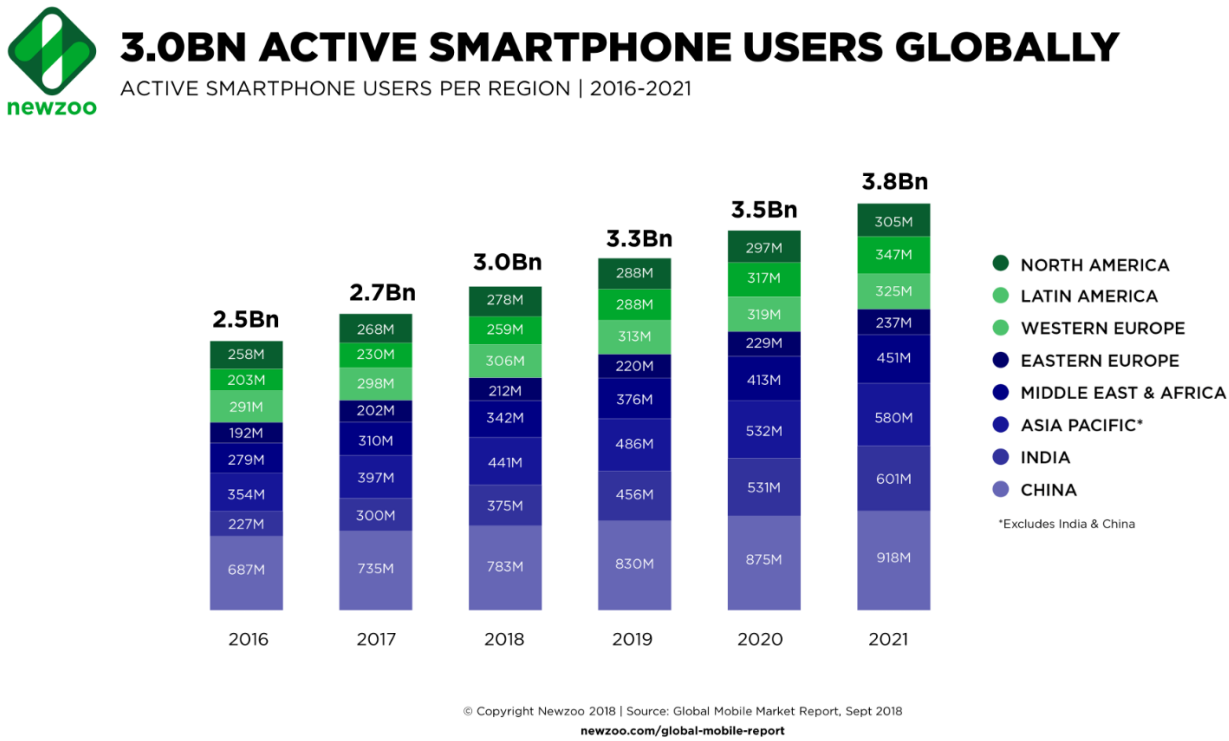


*Figure 12: Smartphone Users Globally (Kooistra, 2018)*

Smartphones use an array of built in sensors which provide data about the location, movement and orientation of the device. It is common for smartphones to have the following sensors built in to provide information on the device's location, orientation and movement (Su, Tong, & Ji, 2014):

**GPS**

Smartphones have built in GPS receivers which enables them to communicate with satellites to determine the location of the device on earth.

**Accelerometer**

Accelerometers detect and record any movement of the device. They can detect the speed the device is moving in any linear direction. In a navigation context accelerometer is used to detect and track the speed of a vehicle.

**Gyroscope**

Gyroscope are sensors which provide precise orientation information about the device, it provides detail regarding the tilt and orientation of the device.

**Magnetometer**

Magnetometers are sensors which can detect magnetic fields, Magnetometers are commonly known as compasses. The sensor data is used by navigation applications to identify the orientation of the device in the context of geographical cardinal directions.

**Pedometer**

Pedometers in conjunction with the accelerometer count the steps and movement of the user. The sensor is widely used by fitness applications for measuring activity.

Smartphones have built in operating systems which allow mobile applications commonly referred to as ‘Apps’ to be downloaded from the operating system market place. The Apps are software applications which are developed specifically for the mobile platform, they perform a specific task or service. Google Maps, the leading navigation application hit 5 billion downloads in early 2019 on the android operating system, highlighting the vast use of navigation applications on smart devices (Rita El, 2019). Downloaded Apps can be given access to the powerful sensor information recorded by the device. As such, third-party app providers can gain access to this sensitive location information about the device position. The power and accuracy of mobile sensors have led to the emergence of Human activity recognition(HAR) systems which are designed to identify the activity currently being undertaken by the user in order to better service the consumer. In terms of navigation applications, activity recognition is heavily linked to form of transportation, a key determining factor for multi modal transportation systems when generating a proposed route. This is discussed in detail in chapter 2.3.3.2.

Privacy concerns have been raised about location tracking on mobile sensors. Huang, Kanhere, and Hu (2010) found that the constant recording of localization sensor data has serious implications for user privacy as the user’s trends and activities can be monitored and combined to unveil the users travel habits and patterns. Privacy threats and privacy protections are discussed in further detail in chapter 2.4 and 2.5.

## 2.3 Multi Modal Transportation Systems

Modern travel planning applications utilise multiple forms of transportation when generating the optimal route for getting to a user's destination. This idea is known as Multi Modal Transportation, a term initially coined by the United Nations in 1980 as the 'the carriage of goods by at least two different modes of transport on the basis of a multimodal transport contract from a place in one country which the goods are taken in charge by the multimodal transport operator to a place designated for delivery situated in a different country'. UnitedNations (1980)

### *2.3.1 Multi Modal Prevalence*

Molin, Mokhtarian, and Kroesen (2016) undertook a study in the Netherlands to better understand travel behaviour and mode of transportation used by commuters in order to develop sustainable travel policies and processes. A key part of the study was to quantify 'the extent to which travellers are multimodal'. Using latent cluster analysis, the researchers identified five single and multi-modal travel groups:

1. Car MM Group

   Mainly used a car for transportation, they had a negative attitude towards cycling and a neutral attitude towards public transportation.

2. Bike MM Group

   Mainly cycled, they had a positive attitude towards public transport and used cars infrequently.

3. Bike + Car Group

   Utilized Bike and Car extensively

4. Car Mostly

   Almost extensively used the car while travelling.

5. The PT MM Group

   Mainly used public transport with cycling also prevalent within the group

The research found that multi-modal transportation was prevalent across the Netherlands and 'all identified clusters are in fact multimodal'.

Satisfaction scores of public transportation systems is another key segment which impacts the use of multimodal transportation systems. Simão (2015) case study studied the effects of multimodal real-time information systems on the travel behaviour of consumers. The

case study did not uncover any global change towards sustainable mobility which is a worrying reveal indicating that the current real-time transportation in place at the time of the case study was insufficient to convert the consumers away from their cars. Abou-Zeid, Witter, Bierlaire, Kaufmann, and Ben-Akiva (2012) found a similar trend when investigating the satisfaction score of Swiss commuters before and after switching to public transportation the results are illustrated below.

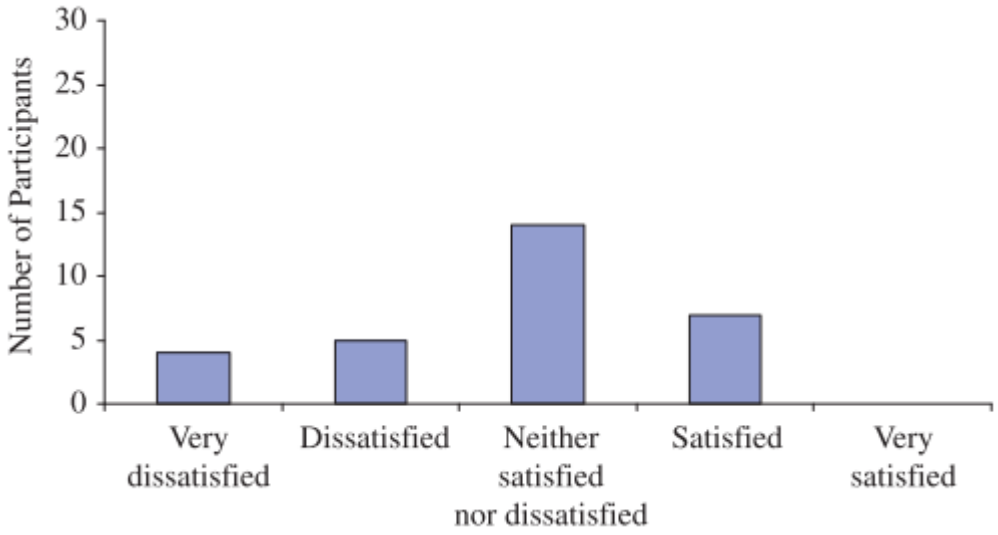


*Figure 13: Post public transport switch(Abou-Zeid et al., 2012)*

Figure 13 illustrates the satisfaction score of car drivers on switching to public transportation. The graph indicates that most car drivers were not satisfied with the switch to using public transportation.

### *2.3.2 Multi Modal System Overview*

To support multi modal initiatives it is critical that navigation planning applications can integrate numerous transportation services into one coherent system which searches for and generates the best route for the customer across different forms of transport, transport operators and transport authorities. Jafri et al. (2013) found that the existing travel planning systems on the market were incapable of aggregating travel related webservices into a single system and proposed the creation of a ‘Smart Travel’ planner which was implemented as an aggregation of serval travel-related APIs. Similarly, Evangelatos et al. (2017) proposed a ‘Super Travel API’ solution to address the problem of aggregating multiple transport vendors into one solution.

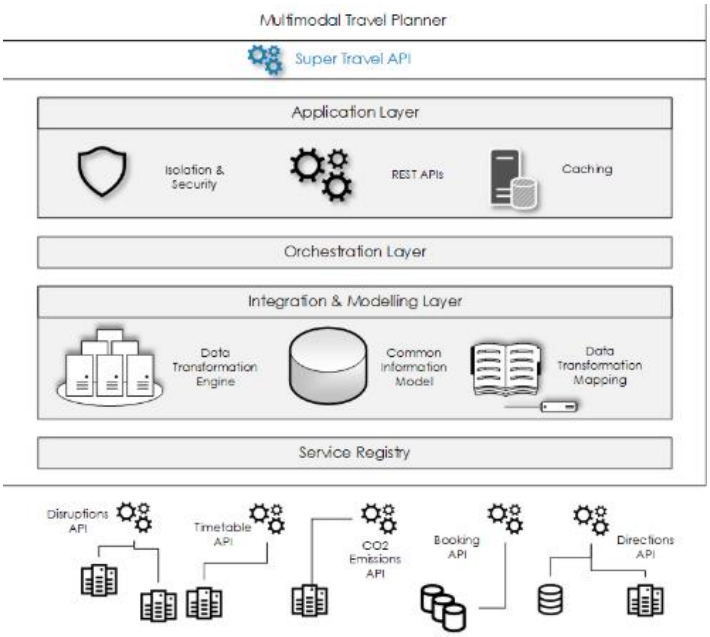


*Figure 14: Super Travel API design (Evangelatos et al., 2017)*

Figure 14 depicts the high-level overview design of the proposed system. The system can generate numerous RESTful web service calls to retrieve the relevant data about services from the relevant authority such as the directions API which will generate the proposed route for the user and the disruptions API which will retrieve information on traffic or delays on the proposed route.

An argument can be made that multi modal transportation systems raise privacy concerns for end users. The system integrates and exchanges location data with third-party services in order to generate the most relevant route for the application. In the multi modal context it is difficult for the end user to understand what parties and services have access to their location data. Damiani and Cuijpers (2013) details that the sharing of location data with third party service providers results in the loss of control of our data which in turns 'translates into loss of privacy'. Privacy protections related to third party integration is further discussed in chapter 2.4.2.

### *2.3.3 Route Generation*

This section is broken into two segments. The first, outlines the shortest-path algorithms used for route generation. The second segment discusses the impact the form of transportation has on route creation.

#### *2.3.3.1 Path generation*

Navigation applications must generate routes to get the user to their destination. Ambrose et al. (2009) discussed the complexity faced by route planners to create the optimal route for each user. The optimal route could mean different things for different people depending on their motivation indicating that the best route 'could be the shortest physical distance, least expensive, least amount of time' or a combination of the three attributes. Routes are generated across a road network which is the representation of Geographical Information Systems spatial data (Karagiorgou & Pfoser, 2012). Road networks are a graph of nodes that represent points of interest on a map. Shortest path algorithms are used to calculate the route between two nodes on the road network, which correspond to the starting and finishing positions for navigation applications.

The two main route planning algorithms are Dijksta's and A* Search. Fan and Shi (2010) described Dijkstra's algorithm as the 'most classical and mature algorithm' for searching for shortest path on a graph. The algorithm computes the shortest path from a single node to all other reachable nodes on a graph. A* Search uses a best-first search shortest path technique which traverses the graph to find the route with the least known heuristic cost. (Chang, Tai, Yeh, Hsieh, & Chang, 2013).

C. L. P. Chen, Zhou, and Zhao (2012) argued that basic approaches for Dijkstra's algorithm and A* Search algorithm are designed for static networks, identifying them as inefficient when calculating and generating more accurate routes to account for real time route information such as traffic data which can affect the optimal route. Similarly, Chang et al. (2013) argue that navigation applications may plan 'erroneous' routes if the route information is out of date.

Several researchers have put forward improvements and enhancements to the above algorithms. Alternative approaches have been used to incorporate traffic variability Ambrose et al. (2009) used a statistical approach to account for traffic variability, the planner generates three different routes, each having a trade-off between expected and variance travel time.

Fan and Shi (2010) put forward techniques for improving the performance of Dijkstra's algorithms which would improve the data storage structure and searching area of the algorithm. Furthermore, Geisberger (2011) indicated that Dijkstra's algorithm is easily extended to include traffic information and public transportation real time information. Chang et al. (2013) proposed a Vehicular Ad-Hoc network-based A* search algorithm that would initially calculate the initial route which is optimal at beginning of the journey. The algorithm would react, and re-plan routes based on real time route information such as

traffic which is exchanged across vehicles to proactively generate an improved route with ‘shortest travelling time and the lowest fuel consumption’.

Xi, Schwiebert, and Shi (2014) argued that origin and destination locations on a route should be private to the user, indicating that this information may be closely tied to the user’s personal life. To ensure privacy the researchers put forward a privacy preserving shortest path algorithm. The solution is based on a principle known as computationally-private information retrieval which involves the requesting of a segment of data in which the requester does not disclose the data used. The application would request all-pairs of shortest path information from the third-party navigation provider. Once retrieved the device itself extracts the optimal route from the dataset not revealing the output of this process to the vendor.

#### *2.3.3.2 Form of Transportation*

The form of transportation is a key element when creating an optimal route to the user’s destination. Shortest-path algorithms use the form of transportation being used to generate appropriate travel time for the route being undertaken. Form of transportation can be provided as an input to many applications when planning the route. However, it is common for navigation applications to determine the form of transportation being undertaken by the user by analysing the trace data of the user’s location provided by the smartphone sensors discussed in chapter 2.2.3. Google Maps provides an Activity Recognition Client API which identifies the form of transportation that is currently being performed by the user (Google, 2018). Google captures and processes GPS location data at intervals passing the data through machine learning techniques and algorithms to identify the most likely activity that is currently being undertaken by the user. The Activity Recognition Client provided by google supports six classifications:

1. IN_VEHICLE
2. ON_FOOT
3. RUNNING
4. WALKING
5. ON_BICYCLE
6. STILL

The activity recognition client enables Google to track the location of the user throughout their entire journey with location trace data shared with Google at regular intervals. The data is used for recalculating travel time and estimated arrival time for the journey. As identified by Xi et al. (2014) the disclosure of location data to navigation applications leads to an increased risk of the user being profiled. Mäenpää, Lobov, and Martinez Lastra

(2017) undertook a study to identify if the form of transportation being used by a user could be identified by using 'Sparse' data, at a rate of tracking the user's location once per minute. The study evaluated the best methods for classifying four distinct forms of transportation: Walk, Bike, Bus and Car using four statistical methods for classification: Bayes Classifier, Random Forest, Neural Network and Autoencoder. The study found that sparse location data was adequate for accurately classifying walking and cycling; however, the four models had similar issues mis-classifying buses and cars. The potential use of Sparse' GPS data would improve the location privacy of consumers reducing the risk of user profiling with a reduced location footprint for the user.

## 2.4 Privacy Threats

This section of the literature review will discuss privacy threats. The chapter will begin by outlining privacy; discussing the definition of what constitutes privacy. The chapter will continue by discussing the emergence of location aware computing and the rise of location-based services. The chapter will conclude by examining the threats to privacy through the exposure of location data to navigation service providers.

### *2.4.1 What is Privacy?*

Brandeis and Warren (1890) originally discussed an individual's right to privacy arguing that the law "*affords a principle which may be invoked to protect the privacy of the individual from invasion*".

In 1948, the United Nations created the *Universal Declaration of Human Rights* which outlined 30 articles affirming to an individual right. Article 12 of the document outlined an individual's right to privacy:

"*No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks*"

(United Nations, 1948)

As previously mentioned, the issue of privacy has evolved in the information technology era. The high availability of sensor-rich devices as discussed in chapter 2.2 has led to an ever-increasing amount of personal data and information about one's position being recorded by navigation service providers. As discussed in the introduction Westin (1967) identified that the key pillar to information privacy is the user. The user must be in control

of their data ensuring that they are aware of “when, how and to what extent about information about them is communicated to others”. The global emergence of location-based services and location aware computing has eroded Westin’s view on information privacy as the user is no longer in control of their data; through the excessive personal data collection practices implied by navigation service providers (Damiani & Cuijpers, 2013). The collection practices of location data by navigation service providers also raises risks of users’ profiling (Freudiger, Shokri, & Hubaux, 2011). Additionally, the storage of location information by third-party service providers raises innate data security threats with data breaches and the leaking of personal information raising additional privacy risks for users (Wheatley, Maillart, & Sornette, 2016).

### *2.4.2 Rise of Location Aware Computing & Navigation Service Providers*

Chapter 2.1 identified that GPS and Network location tracking technologies innately protect the location of the user. The radio frequency waves are unidirectional meaning that the receivers do not relay the user location back to the transmitters (L. Chen et al., 2017). Similarly, as discussed in the navigation applications section the privacy of the user’s location is preserved when using Portable Navigation devices and In-Dash navigation devices. The reason for this is that the location of the user and route information is generated through a built-in algorithm on the device itself. Therefore, the key threat to consumer privacy in the navigation applications market is the use of third-party navigation service providers which calculate the route for the user based on their origin and destination locations. Third-party applications generate the route information for the user off the system which results in the loss of control of their location data (Damiani & Cuijpers, 2013). Third Party providers are continually striving to provide the best service for their customers that will give them a competitive advantage in their market. Likewise, consumers wish to use the best solution which will fulfil their need. In the navigation applications market, organisations are capturing and processing colossal volumes of user’s location data to gain insights into their journey and other users journeys to improve their quality of service, for example providing real time traffic information. As already highlighted the staggering volume of data can be visualized through Google’s ‘Timelines’ product. This has led to location centric services being created commonly known as location aware computing or location-based services. Location aware computing raises privacy concerns for users. Beresford and Stajano (2003) discussed the implications of location aware applications which track user movements. The researchers found that location aware applications collected ‘enormous amounts of potentially sensitive information’ which had implications to the user’s location privacy. Similarly, Barkhuus and Dey (2003) found that location based services raised more concern for users than position

based services, which uses the device location to determine the position while location-based services transfer location information to third-parties. Location based services (LBS) offer users superior services and accessibility to applications which would not be possible without the knowledge of the consumers location. However, the dilemma for the user is that more personal location information is shared with the applications; As the quality of the service improves the privacy risks associated with the use of the service increases (Xu, Teo, Tan, & Agarwal, 2009).

The World Economic Forum (2017) created a framework called Valuing Person Data which is illustrated in figure 15 below. The framework was created to provide a taxonomy for the types of personal data captured through the consumption of digital platforms and services. The framework investigates the commercial side of personal data capturing, identifying that organisations capture data of the user and combine their information to gain insights into the consumer behaviour and to build a profile of each user. The third step of the framework 'Digital Footprints' explicitly mentions location data for the creation of the footprint for the user. Through the creation of a digital footprint, organisations are capable of understanding traits, tendencies and locations of the consumer allowing them to target specific services and advertisements at the user.

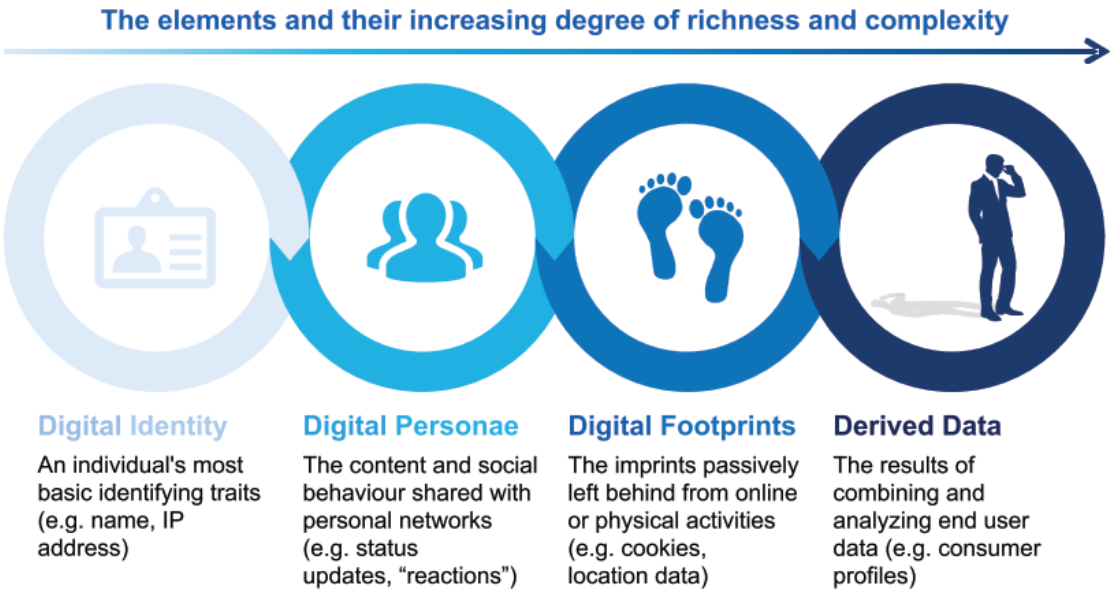


*Figure 15: Valuing Person Data Framework (World Economic Forum, 2017)*

Damiani and Cuijpers (2013) suggested that the interaction between the service provider and the user's device should be minimized. The amount of information sent to the service provider should be the minimal amount to provide the service. Similarly, Mateosian (2013)

put forward that the data handling practices must align with an organisation's values. Organisations should not simply track the user because they are capable of it; they should only capture the data they require to provide an adequate service for the consumer.

### *2.4.3 Risk of User Profiling*

The World Economic Forum (2017) identified digital identity as the initial element of the valuing person data framework, Similarly, Barkhuus and Dey (2003) argued that the core principle to privacy is 'Identity', with location information one of the key attributes to a person's Identity.

As outlined in the valuing personal data framework (Word Economic Forum, 2017) the creation of a profile for a user has commercial benefits for organisations allowing them to target specific services and advertisements. Hasan, Habegger, Brunie, Bennani, and Damiani (2013) discussed the privacy concerns raised by the creation of user profiles. The researchers identified that location data was one attribute of the user profile created by organisations. Leontiadis, Efstratiou, Picone, and Mascolo (2012) discussed the balancing act between mobile privacy and supporting the mobile application market which depends on advertising. The mobile applications market as discussed in chapter 2.2.3 allows users to download a multitude of mobile applications to their device that will provide a task or service to the user. Leontiadis et al. (2012) identified that the smartphone mobile marketplace was driven by advertising with many of the applications requesting location data and other personal information to gain insights into the consumer. Through the collection and analysing of personal data amassed by third-party applications there is an increased risk of the user being identified and profiled. Bettini, Wang, and Jajodia (2005) discussed the idea of '*Service Request Linkability'* which is a measure of the possibility of linking two anonymous requests back to the same user who issued the requests by analysing a set of requests issued to a service provider. The increased *Linkability* of service request poses a threat to user privacy as a single user's service requests can be identified increasing the risk of profiling the user. There is a lot of research done on the topic of user profiling. Dalenius (1986) initially discussed the topic of quasi-identifiers, which are segments of personal data stored which alone are not unique identifies but which combined and correlated with additional quasi-identifiers could result in a unique identifier being revealed which results in the user being profiled. Freudiger et al. (2011) further discussed *location-based Quasi-identifiers* which are traces of anonymous location data which combined and correlated can significantly increase the risk of the user being profiled. The study identified that the correlation of travel patterns has a high risk of unveiling the user's home and work locations. Similarly, Vicente, Freni, Bettini, and Jensen (2011) identified that through the access to location and time information

organisations could identify sensitive information about the consumer; citing personal revelations such as health problems or habits and tendencies. Freudiger et al. (2011) discussed the concept of a 'Threat Model' for location-based services (LBS). The model discusses the threat to privacy posed by the data collection practices of location service providers who actively monitor and record user's location. The researched considered that the goal of LBS is to reveal the true identity of the user to understand their traits and points of interests. This raises grave concerns for information privacy as through the collection and monitoring of user's location over a set of days, the application can infer the activities and key points of interests to the user which may unveil their true identity. The Threat model is illustrated below in figure 16.

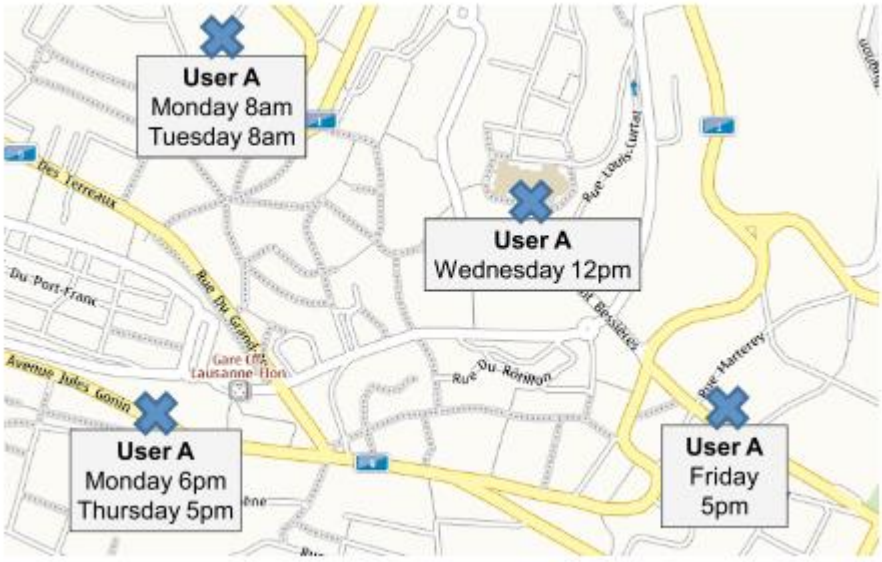


*Figure 16: Threat Model (Freudiger et al., 2011)*

### *2.4.4 Data Breaches*

Spiekermann-Hoff (2012) identified personal data as the "the asset at the heart of many companies' business models", Organisations are massing huge amounts of personal information on consumers resulting in the users' loss of control over their personal data (Damiani & Cuijpers, 2013). The loss of control of personal data erodes personal privacy, Wheatley et al. (2016) identified data breaches as an 'Extreme Risk' to user's privacy. The researchers proposed that the capture and storage of vast quantities of personal information has increased the privacy risk associated with data breaches. (Stevens, 2012)

defined data breaches as a security incident which results in a loss, theft or the unauthorized access to personal information. Similarly, the United States Department of Health and Human services categorized data breaches as a security violations in which confidential or protected information is “copied, transmitted, viewed, stolen or used by an unauthorized member” (US Department of Health and Human Services, 2015). The frequency and scale of data breaches is alarming. The Privacy Rights Clearinghouse, a website which tracks data breaches found that as of April 2019 11,575,804,706 records have been breached from 8,804 breaches which have been publicised since 2005 (Privacy Rights Clearinghouse, 2019). Cisco (2016) forecast that the volume of data traffic will increase seven-fold between 2016-2021 with such vast increases in data volume the privacy threat posed by data breaches becomes greater (Wheatley et al., 2016). Cisco’s forecast is illustrated in figure 17.

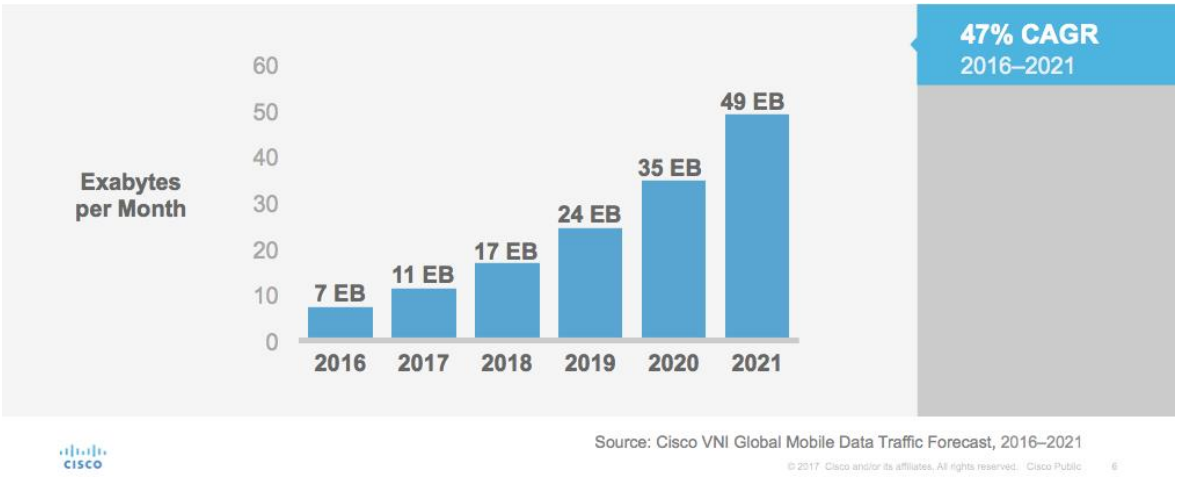


*Figure 17: Mobile Data Traffic Forecast 2016-2021 (Cisco, 2016)*

Wheatley et al. (2016) discussed the large-scale sharing of breached data in underground markets. The sharing of breached data increases the privacy risk for users, there is a far greater chance of the user being profiled and identified with multiple unauthorized parties gaining access to the data from the breach. As discussed in the user profiling section, access to personal information including location data allows the data observer to identify the user through analysing location-based Quasi-identifiers (Freudiger et al., 2011). Data breaches do not just affect consumer privacy they also have substantial costs on the organisation affected. The IBM Security Institute study on the cost of data breaches for organisations found that the average cost for each breached data record was $148 in 2018 rising from $141 in 2017 (IBM, 2018).

### *2.4.5 Location Privacy Threats*

Krumm (2009) survey on location privacy identified three preferences of people when disclosing their location information:

1. When

   People are more likely to reveal historic location information rather than current or future location.

2. How

   People are more comfortable revealing location information to people they know and trust rather than strangers.

3. Extent

   People are more comfortable reporting ambiguous location rather than providing their precise location.

The three main attributes to the location privacy preferences captured by Krumm can be highlighted in the key threats already identified in this section. The vast personal data collection practices and creation of user profiles by organisations impacts all three; 'When', 'How' and 'Extent'. The location data captured by organisations is continuous, extension and is captured by strangers. Similarly, the data breaches impact all three with data breaches revealing comprehensive and sensitive information to unauthorized personal. Additionally, to the threats discussed there are also location specific attacks that raise privacy concerns for the user. Through the capture of vast quantities of location data people are vulnerable to location privacy attacks classified as 'Attack Knowledge' attacks by Wernke, Skvortsov, Dürr, and Rothermel (2014). Wernke et al. (2014) spilt Attacker Knowledge attacks into two levels based on the attackers acquired information on their target's location – the user. The attacks are classed as temporal information attacks or context information attacks. In temporal information attacks, the attacker has basic current and historic location information of the user with little or no context linking the data points together. In context information attacks the attacker has a far greater understanding of the user's behaviour through additional context information such as time and date along with the location data points. Context information attacks allow users to build a user profile based on their movements, Similar to the 'Threat Model' created by Freudiger et al. (2011) which builds a profile of the target through tracking of their movements over a period of time. Attacker knowledge attacks exploit users through knowledge of location information. One suspected location privacy attack as identified by Vicente et al. (2011) is 'Absence Privacy'. Absence privacy relates to the revealing to adversaries information that the user isn't at a certain location. The concept of 'Absence Privacy' suggests that through the

exposure of location information personal risks are heightened. The paper cites the example of burglary based on knowing that the owner is not present at their property, criminals may plot to steal their belongings. Attacker knowledge attacks are illustrated in figure 18.

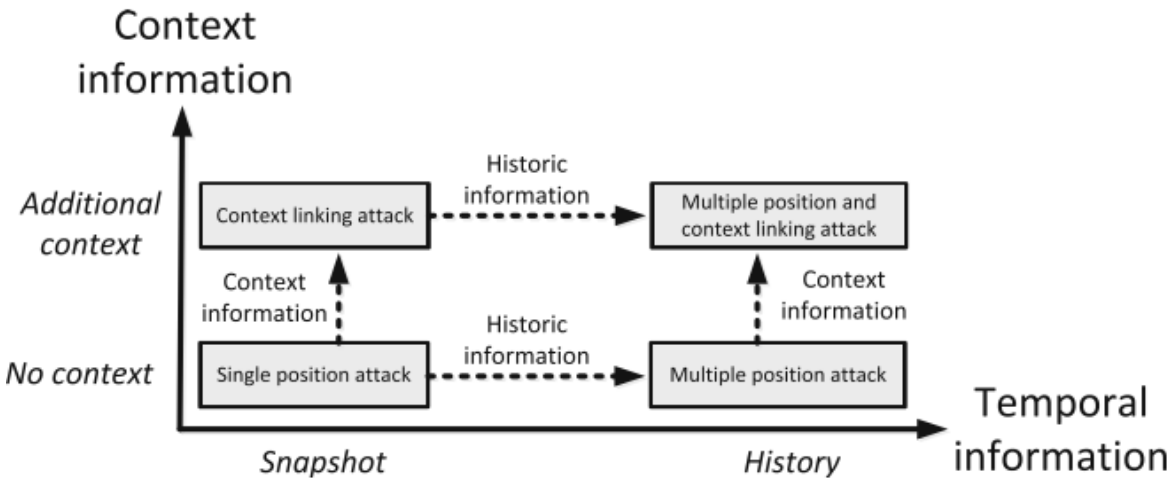


*Figure 18: Attacker Knowledge Attacks (Wernke et al., 2014)*

## 2.5 Privacy Protections

This section of the literature review will discuss the protection of privacy. The legislative protections in place to protect privacy will be outlined. The technological protections used to preserve privacy will be discussed and it will be followed by an investigation into location privacy preserving techniques used to protect users' location information.

### *2.5.1 Legislation Protections*

Legislation protections refer to any laws which protects users privacy. Legislation privacy protections is not a new concept, Brandeis and Warren (1890) initially discussed a person's legal right to privacy under the United States constitution, the United Nations (1948) further defined privacy as a human right. With the rise of personal data capturing it is imperative for the privacy of consumers that strict data privacy and data processing laws are in place to protect user's personal data. Alepis, Politou, and Patsakis (2018) identified that the rise of ubiquitous computing and the emergence of big data as a commercial model have 'severe implications' to human civil rights, specifically to people's rights to privacy and data protection laws. Similarly, Wachter (2018) discussed the undermining of privacy protection laws by organisations through their personal data

collection practices which "collect, share, and store large and varied types of personal data".

One of the leading legislation frameworks is the General Data Protection Regulation (GDPR) created by the European union. Buttarelli (2016) identified GDPR as the 'digital gold standard' for data protection laws in the world. Similarly, Goddard (2017) identified GDPR as having a 'Global Impact' on privacy legislation around the world citing that GDPR emphasises privacy as a fundamental human right. GDPR came into law under the European Union protection framework on the 25th of May 2018. GDPR encapsulates all legal regulation regarding the capturing and processing of data. GDPR "emphasises transparency, security and accountability by data controllers and processors, while at the same time standardising and strengthening the right of European citizens to data privacy" (Data Protection Commision, 2018). GDPR regulation ensures that organisations who utilise European Citizens data must do so in a legal, transparent manner using an accepted security protocol. Goddard (2017) identified the six core privacy protection principles behind GDPR:

1. Fairness and Lawfulness
   Personal data capture must be done in a legal manner with the awareness and consent of the data subject.

2. Purpose Limitation
   The data captured by an organisation must be done for a specific purpose, the purpose and use of the data must be clearly stated by the organisation. As discussed by Damiani and Cuijpers (2013) organisations must only capture the data for the purpose of providing the service to the user. The data should not be used for any other endeavour.

3. Data Minimisation
   Service Providers must only capture the data needed to provide the service. Organisations have the responsibility of minimising the volume of personal data captured, they must only capture data that is necessary to provide the service (Mateosian, 2013).

4. Accuracy
   GDPR outlines that personal data stored on a data subject must be accurate and where necessary kept up to date.

5. Storage Limitation
   Personal data must only be stored for the time “necessary for the purposes for which the personal data are processed” (General Data Protection Regulation, 2018). The time that the personal data is stored should clearly be outlined by the service provider.

6. Integrity and Confidentiality
   Data should be stored and processed in a secure manner to protect against unauthorised processing. The necessary data encryption and obfuscation technologies and techniques should be used to protect the data (Papageorgiou et al., 2018).

General Data Protection Regulation ensures organisations who capture and process personal data are held accountable for their data management practices. Organisations who fail to comply with the regulation are heavily fined. For serious breaches of the regulation Organisations can be fined for up to 4% of their worldwide revenue or twenty million euro, depending on which figure is higher (Tankard, 2016). GDPR is an EU Directive for all member states, requiring all members to write the regulation into law (Goddard, 2017). As noted by Alepis et al. (2018), the use of an EU directive signals a turning point for the European Union with all member states protected under an EU level law rather than individual member state law. Similarly, Albrecht (2016) emphasised the legal certainty and unity GDPR brings to the personal data processing laws for European Citizens through unified regulation across all countries within the European Union.

General Data Protection Regulation puts the control of personal data back into the hands of the data subject rather than the data controller with the inclusion of a number of rights to the regulation (Alepis et al., 2018) namely:

1. The right to erasure
   Commonly known as the ‘Right to be forgotten’ allows data subjects to request service providers to erase all personal data stored on them.
   The ‘Right to be Forgotten’ term emerged in May 2014 when the European Court of justice ruled that European Citizens had the right to request search engines to remove results that were linked to them (Newman, 2015).

2. The right to revoke consent
   Consent by definition according to Oxford Dictionary (2019a) is the

"Permission for something to happen or agreement to do something". In the context of personal data, consent is used to provide data controllers and collectors with legitimate grounds for data collection and processing. The revocation or withdrawal of consent ceases the collection and processing of personal information by Organisation for that particular user (Alepis et al., 2018).

3. The right to Portability
   Commonly known as the 'Right to Access' allows data subjects to request and obtain any personal information stored on them by data collectors. De Hert, Papakonstantinou, Malgieri, Beslay, and Sanchez (2018) argued that the right to portability embedded into GDPR is a large development towards user-centric privacy allowing users to view, evaluate and ultimately control the personal data stored on them by data collectors and organisations.

The European Union's GDPR regulation is leading the world in providing European Citizens a robust, strict and accountable framework to govern the personal data processing practices used by Organisations. Other Policy makers and governments are adjusting their privacy policies to better protect their citizen's personal data (Albrecht, 2016). The United Nations (2018) build on GDPR to define ten principles to govern the use of personal data by United Nations organisations. The purpose of the principles as outlined by the United Nations (2018) is to: Standardize the protection policies of organisations, facilitate accountable personal data processing and to ensure human rights of consumers are maintained. Brazil introduced Generate Data Protection law (GDPL) in 2018 to improve the controls users have on their personal data (Wilkinson, 2018). Similarly, in Japan the 'Act on the Protection of Personal Information' was amended in 2017 to better protect personal data and people's right to their data (Personal Protection Commission, 2017). For other countries including the United States and Canada, Europe's GDPR stands in stark contrast with their weak privacy policies (Newman, 2015).

### *2.5.2 Privacy Policy*

As previously discussed, Privacy policies are legal documents which outline the data handling practices employed by organisations. Westin (1967) identified that the goal of privacy policies is to reduce fear that users information will be disclosed and to increase the trust between the consumer and the organisation in terms of their data processing management. Similarly, Besmer, Watson, and Lipford (2010) emphasised the need for consumers to understand the information stored on themselves, which builds a mutually

beneficial relationship in which the consumer has control over their data and businesses can ethically use their information to improve their services. Wu, Huang, Yen, and Popova (2012) suggested that privacy policies are built around five principles for fair information gathering practices:

1. Notice
   Consumers should be notified of the data collection and processing practices prior to the any data being captured.

2. Choice
   Consumers should have a choice about what personal data is collected on them and have control over what data is stored on them.

3. Access
   Consumers should have access to the data stored on them. Allowing the consumer to evaluate the accuracy and completeness of such data. As previously highlighted GDPR has access rights embedded into law with the 'Right to Portability' (De Hert et al., 2018).

4. Security
   Consumers data should be stored in a secure manner with Anonymisation and Pseudonymization processes done on the dataset to ensure that the data is unidentifiable (Esayas, 2015).

5. Enforcement
   Enforcement refers to the legislation protections in place to ensure that the data collection practices of organisations is kept to a high standard. GDPR as previously discussed is the leading legislation in the world for protecting consumer privacy through the strict enforcement of the legislation with large fines for non-compliance (Tankard, 2016).

Privacy policies make consumers aware of the data capturing and processing practices of an organisation or service. Many service providers require the user to agree to the privacy policy before commencing their use of the service. However, it could be argued that privacy policies are not fit for purpose, Wachter (2018) identified that privacy policies often fail to clearly communicate the impact of data processing and capture policies have on the user's privacy. A 2017 Deloitte survey of 2,000 participants found that a whopping 91% of consumers do not read the privacy policy before agreeing to use the service (Cakebread,

2017). Privacy policies are notoriously long and difficult to understand, Schwab (2018) found that the average length of privacy policies for the top twenty mobile applications is 3,964 words. Included in the top 20 was Google maps, the largest navigation application on the market with a 4,528-word privacy policy.

### *2.5.3 De-Identification of Personal Data*

As outlined in the navigation application section, navigation applications capture and process large volumes of personal information. The personal data must be stored in a format that makes it unidentifiable, making it impossible to link data records back to an individual. Barkhuus and Dey (2003) argued that 'Identity' is the core principle behind privacy, if a person can be identified their privacy is exposed. Freudiger et al. (2011) highlighted that through the use of location based quasi- identifiers the risk to users being identified was greatly increased. To combat the risk of user profiling and identification one of the core privacy protection measures is the de-identification of personal data. The de-identification process is commonly known as data anonymization (Esayas, 2015). Zhang, Yang, Liu, and Chen (2014) defined data anonymization as the process of hiding the identify and/or sensitive data of data subjects. Through the use of data anonymization the privacy of the individual can be maintained while data processing organisations can still use the data for data mining and analysis (Zhang et al., 2014). The General Data Protection Regulation (GDPR) requires all data controllers to run de-identification procedures over their datasets to ensure data records cannot be linked back to individuals, (Esayas, 2015) identified that de-identification according to GDPR contains two core processes, data anonymization and pseudonymisation.

Bettini et al. (2005) discussed data anonymization techniques for protecting users privacy against location-based identification. The researchers proposed an anonymous location-based service model that would isolate the service provider and the users through the use of a trusted service. The trusted service provider would use a privacy preserving framework based on K-Anonymity. Huang et al. (2010) defined k-anonymity as a dataset which is indistinguishable with respect to some chosen attribute. K-Anonymity ensures that common identifying information is obfuscated.

Kapadia, Triandopoulos, Cornelius, Peebles, and Kotz (2008) proposed tessellation, a K-Anonymous based location privacy preserving technique. Tessellation provides privacy through the generalisation of location data; location granularity is greatly reduced. A simple explanation of tessellation was discussed by Huang et al. (2010) who cited the example of replacing a street-level location data with a city-level equivalent. Huang et al.

(2010) argues that tessellation is unsuitable for applications which require concise location accuracy such as navigation applications, as the location precision is greatly reduced.

To produce K-Anonymous data across large datasets significant computational power is required. To overcome the performance issues with large scale datasets Zhang et al. (2014) proposed a the use of MapReduce jobs to optimize and improve performance. Wachter (2018) identified some weaknesses of anonymisation to prevent profiling arguing that it is possible to reverse engineer anonymisation algorithms to retrieve the personal information of users. The researched noted that data cannot be completely anonymised without removing its analytical value.

The second core topic of de-identification as identified by Esayas (2015) is pseudonymisation. Esayas (2015) argues that pseudonymised datasets differ from anonymised datasets; pseudonymised datasets allow an individual data subject to be singled out and linked across different datasets. While, in anonymised datasets it is impossible to single out a data subject. Pseudonymisation is the process of making a data record less identifiable while maintaining its suitability for data mining and legislation compliance. As discussed in the legislations protection section one of the core rights invoked by GDPR is the right to portability, in a Pseudonymised dataset all the data stored on a single user is retrievable. Many researchers have identified encryption as one of the key processes involved in Pseudonymization (Esayas, 2015) (Nikolaenko et al., 2013). Tyagi and Sreenath (2015) identified cryptography-based approaches as a core location privacy technique when sharing location information with applications. Through the use of encryption, the privacy of the user is maintained, only the trusted application with the shared secret key is capable of viewing the location data. Tyagi and Sreenath (2015) state that the approach is robust with user privacy only being breached if both the authenticator and the service provider are compromised. Similarly, L. Chen et al. (2017) state that through the use of encryption the personal information shared across a network is kept secure and confidential. Esayas (2015) noted that encryption as a privacy-enhancing measure is only as strong as the encryption algorithm and encryption key used. In Tankard (2016) study on organisations focus for achieving GDPR compliance, Encryption was identified as the number one priority to achieve GDPR compliance. The survey result is illustrated in figure 19 below.

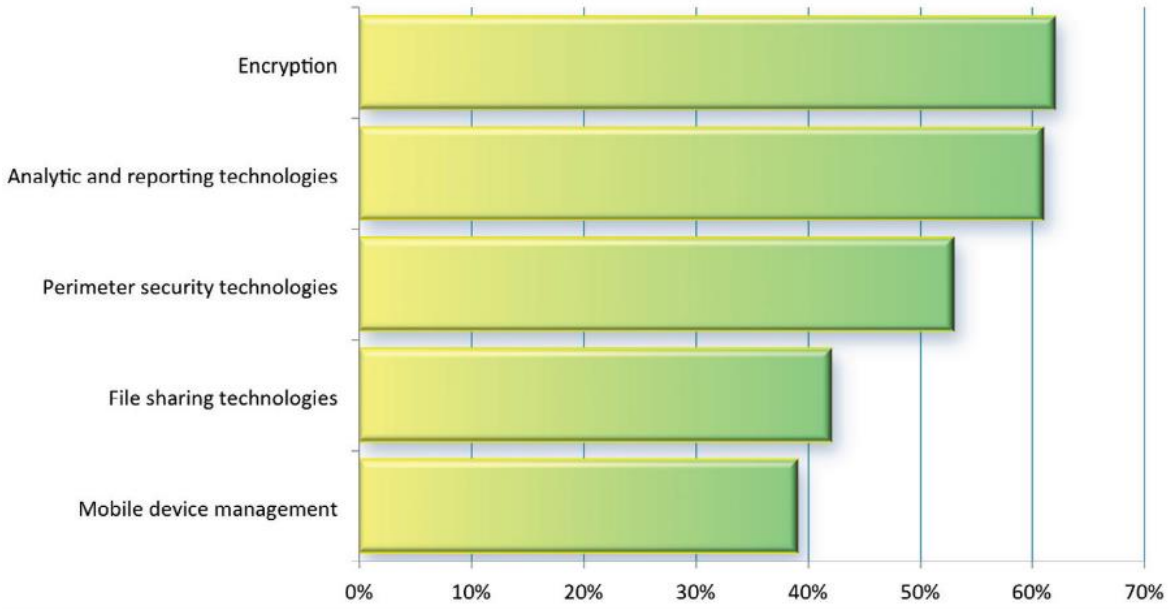


*Figure 19: Investment for achieving data protection compliance (Tankard, 2016)*

### *2.5.4 Mobile Application Permissions*

As discussed in the Navigation Device section, Smartphones are the most widely used device for navigation applications. Smartphones are sensor rich devices capable of tracking movement and the position of the user based on embedded sensors (Wan et al., 2011). As previously discussed, Smartphones provide a market place to download third-party mobile apps which can gain access to these sensors and thus allow them to track and record the location of the user. This raises privacy concerns for users with the loss of control of their data. One of key protections in place for smartphones is the permissions framework. (Lin et al., 2012) identified that the smartphone permissions framework is designed to protect users in two ways:

1. Limit mobile apps access to sensitive resources
   Third-Party applications by default will have no access to sensitive user information and sensors. The permission framework ensures that third party applications will be unable to use the resource unless consented by the user.

2. Only allow application permission to resources that are agreed to upon installation.
   The user must explicitly consent to what resources and sensors the newly downloaded application can access. The permission framework enables smartphone users to take control of what data applications can access.

Lin, Liu, Sadeh, and Hong (2014) identified that many third-party libraries bundled inside downloaded applications request resources and permissions that are not associated to the apps purpose. The researchers cited the example of location data being requested by an application only because a 3rd party library bundled inside the app requires the permission. The nine categories of 3rd-party libraries as identified by Lin et al. (2014) can be found in appendix 2. Chitkara, Gothoskar, Harish, Hong, and Agarwal (2017) undertook a study to identify the extent applications request permissions that are not core to their functionality. The researchers found that a large proportion of applications requested permissions that had no effect on their service offering. Furthermore, the researchers found that many of these applications shared the users data with third-party analytics providers without the knowledge and consent of the user.

Almuhimedi et al. (2015) discussed the implications of permission sharing with third party applications citing that smartphone users are regularly unaware of the data collected on them by apps. The researchers undertook a study to evaluate the benefit of a privacy notification nudges, the notification would integrate with the permission manager of the device to identify when and what permissions applications are requesting. The study found that 95% of the participants reviewed their application permissions within one week of the study commencing, indicating the need for more oversight on the volume and type of information being recorded by third party applications on smartphones. A screenshot of the privacy nudges can be seen in figure 20.

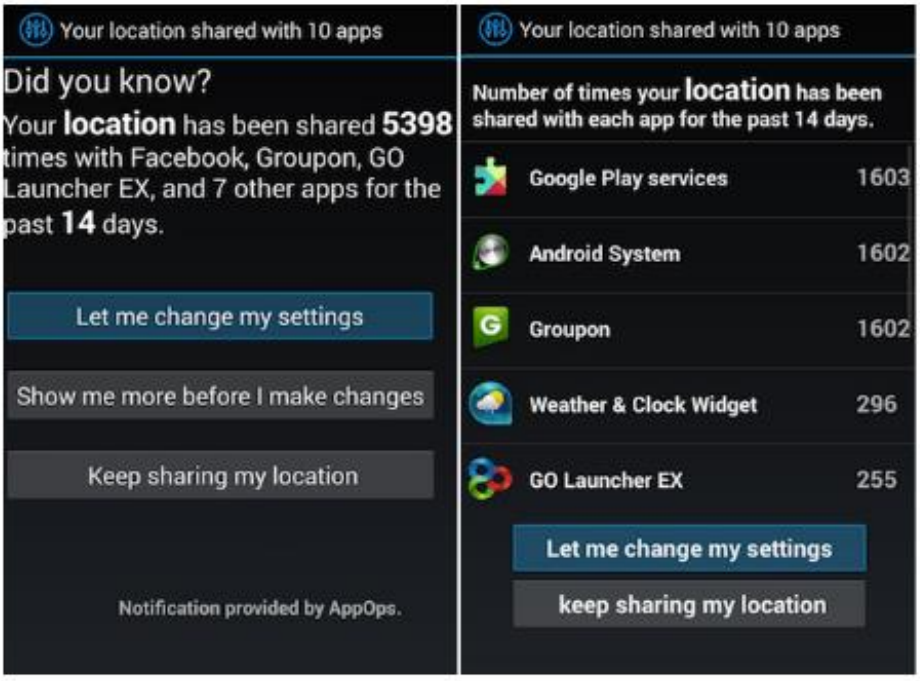


*Figure 20: Privacy Nudge Screenshot (Almuhimedi et al., 2015)*

### *2.5.5 Location Privacy Preserving Mechanisms*

Shokri, Theodorakopoulos, Troncoso, Hubaux, and Le Boudec (2012) identified Location privacy preserving mechanisms as having two specific focuses:

1. The accuracy of recovering actual user location information from anonymous location traces
2. Creating an optimal privacy metric for measuring the effectiveness of privacy preserving mechanisms.

Location-Based Quasi-identifiers as discussed by Freudiger et al. (2011) raise user location profiling and identification risks through analysing anonymous location trace data. The better the location preserving mechanism used the lower chance an adversary can identify the user's true location.

Kido, Yanagisawa, and Satoh (2005) proposed the 'position dummies' technique to improve location privacy; the solution generates and sends several fake position data points to the service provider along with the true position of the user. All location queries for both the dummy position and the real position of the user is processed making them indistinguishable from one another, there is no way for the service provider to know which is the real position thus the location privacy of the user is protected (Dürr, Skvortsov, & Rothermel, 2011) (Shokri et al., 2012). One key advantage for position dummies identified by Wernke et al. (2014) is that the protection is built into the device itself, the device can generate the location dummies without requiring a location service provider. Figure 20 illustrates the dummy concept with the dark nodes representing real user location and the light nodes presenting dummy (fake) location data.

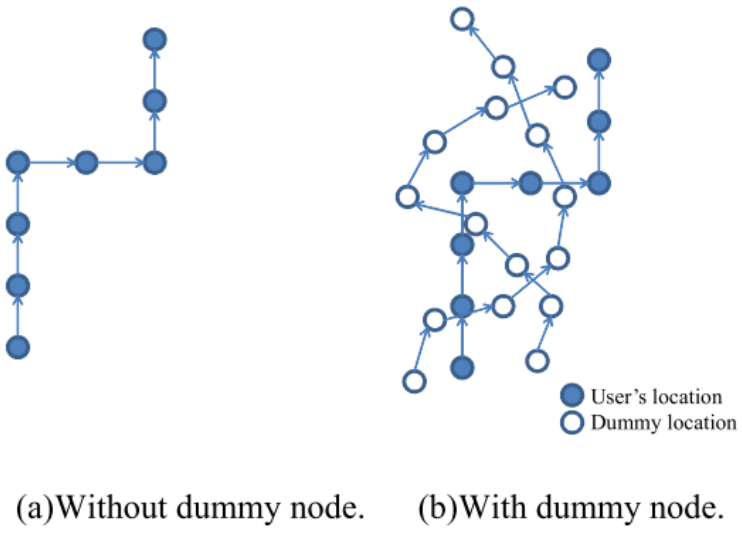


*Figure 21: Position Dummies Concept (Kasori & Sato, 2015)*

Beresford and Stajano (2004) proposed the 'Mix Zone' technique to protect location privacy. The mix zone model creates geographical boundaries in which users location information cannot be recorded. If the data subject enters a 'mix zone' their pseudonyms is changed to protect their identity (Wernke et al., 2014). The 'mix zone' technique uses a trusted middleware service provider as an interface between the underlying location tracking device and untrusted third-party applications, the user will manage and configure their zones through the middleware application which will then manage and change the users pseudonyms to maintain the user's their privacy within these zones. Similar to the mix zone concept, Kasori and Sato (2015) discussed the 'cloaking region' as a technique to protect location privacy. In the cloaking region model the location of the user is anonymized in a designated spatial region divided into specific grids. The level of location privacy is based on the number of spatial grids in which the users location is blurred. The grid layout in figure 22 below illustrate the cloaking region concept.

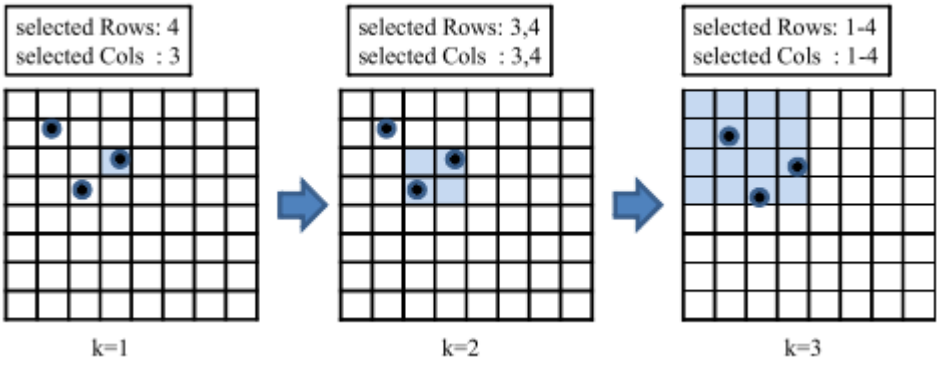


*Figure 22: Cloaking Region (Kasori & Sato, 2015)*

# 3 Research Methodology

This chapter discusses the research methodology used throughout this paper. The chapter will discuss the importance of research design, the chosen methodology and the steps involved in the execution of the methodology. The chapter concludes by discussing the lessons learnt and limitations of the chosen methodology.

## 3.1 Research Design

The (Oxford Dictionary, 2019) defines research as 'The systematic investigation into and study of materials and sources in order to establish facts and reach new conclusions'. Research is done to uncover knowledge in a systematic way, increasing the expertise in a specific area (Saunders, Lewis, & Thornhill, 2007). The use of a systematic approach ensures that the correct procedures are in place to identify, examine, criticise and evaluate literature in a reliable repeatable manner (Saunders et al., 2007). The systematic approach is made possible through the correct research design which provides a framework and plan for undertaking a research endeavour. The correct research design provides a blueprint on how the research can be carried out in order to answer the research objectives outlined in a study (Creswell & Creswell, 2017). In this research a 'Systematic Literature Review' research design was used to answer the research objectives. The methodology is described in detail in the subsequent chapter; 'Systematic Literature Review'.

## 3.2 Systematic Literature Review

Following on from the research design discussion which requires a systematic approach to research, it was decided to use a Systematic Literature Review as the methodology for this paper. Okoli and Schabram (2010) identified systematic literature review as a 'free-standing article whose specific and entire focus is to review research on a subject'. Okoli and Schabram (2010) indicate that a standalone systematic literature review is in itself a complete research output. Systematic literature review originated in the health science field with Fink (2005) defining a systematic literature review as '"a systematic, explicit, comprehensive and reproducible method for identifying, evaluating and synthesizing the existing body of completed and recorded work produced by researchers, scholars and practitioners'. Fink (2005), proposes that a standalone literature review may be done to understand the knowledge within a professional practice, In the case of this project the methodology was chosen to complete a comprehensive literature review of the impact navigation applications have on the location privacy of users. There has been no such

research published that provides a complete overview of the location tracking technologies, devices and privacy threats and protections in the navigation applications market.

Okoli and Schabram (2010) built on Fink (2005) standalone literature review to provide a guidelines for completing an information systems literature review. Okoli and Schabram (2010) emphasised that a different approach was needed in an information systems literature review process as opposed to the health sciences, as information systems 'is a combination of social science, business, and computing science'. Okoli (2015) building on his work of 2010 proposed an eight-step guide to completing a rigorous systematic literature review:

1. Identify the purpose of the literature review
2. Define research question and draft protocol
3. Search for Literature
4. Practical screening for inclusion
5. Quality appraisal for exclusion
6. Data extraction
7. Synthesize studies
8. Write the review

This research uses Okoli (2015) eight step approach as a guide. The subsequent sections will discuss the processes and procedures used for each step. The systematic literature review process is illustrated in the appendix 3.

## 3.3 Purpose of Literature Review

The research title was chosen to investigate the navigation application industry which has been transformed through the emergence of location tracking technologies and the rise of ubiquitous computing devices capable of continuously tracking precise user movement and location. The purpose of the literature review is to gain an understanding of how modern navigation applications work by investigating the underlying technologies and devices used by navigation applications to better understand the impact such applications have on the consumers' location privacy. The research examines the two key dimensions of location privacy: Privacy Threats posed by the use of navigation application and Privacy Protections in place to protect the privacy of the consumer.

Ultimately, the purpose of the literature review is to answer and provide an understanding around the three research objectives outlined in chapter 1.3:

1. Gain an understanding of how navigation applications work
2. Identify the potential privacy threats imposed through the use of navigation applications
3. Identify the privacy protections in place to protect users of navigation applications

The target audience for the research as discussed in section 1.5 are privacy regulators and policy makers. Additionally, the research is targeted at privacy aware consumers who are curious to understand how their privacy may be impacted through the use of navigation applications.

## 3.4 Research Question and Protocol

*An Investigation into location privacy of Navigation Applications: A Systematic Literature Review*

Okoli (2015) outlined that a formulated research question should be a short statement which conveys three characteristics of the research:

1. Target Audience
2. Purpose
3. End Use

The research question outlined in this paper accurately conveys all three characteristics. The title includes the two key themes behind the research; navigation applications and location privacy, ensuring that the purpose is clear to the reader. Similarly, privacy regulators and curious navigation application users can easily identify the contents of the paper through reading the title. Through the inclusion of the words ‘A Systematic Literature Review’ in the title of this paper the reader can identify the specific type of research completed.

The second step of this section is to identify the research protocol. As discussed in this chapter, the protocol for this paper is the eight step approach to writing systematic literature review in information systems as purposed by Okoli (2015).

## 3.5 Literature Search Process

This section will outline the search approach used to identify the relevant papers related to the study. The study consisted of five core topics:

1. Navigation Applications
2. Location Tracking Technologies
3. Multi-Modal Transportation
4. Location Privacy Threats
5. Location Privacy Protections

Google Scholar was the sole search engine used through the literature search process. Table 1 below outlines the search terms used to generate relevant papers along with the number of hits for each search. The use of quotes greatly improved the accuracy of the results and improved their relevance to the topic which is highlighted by the reduced count for quotation searches. Quotation searches are direct string searches which will only return results that have that exact string. Furthermore, the use of combined direct string searches through the use of the keyword 'AND' greatly reduced the count of articles returned improving their relevance to the search.

| Search Term | Count |
|---|---|
| Navigation Applications | 3,240,000 |
| "Navigation Applications" | 19,200 |
| Navigation App | 419,000 |
| "Navigation App" | 2,090 |
| Location Tracking | 3,160,000 |
| "Location Tracking" | 90,400 |
| Location Tracking Technology | 3,260,000 |
| "Location Tracking Technology" | 1,370 |
| Multi Modal Transportation | 206,000 |
| "Multi Modal Transportation" | 7,600 |
| Co Modal Transportation | 138,000 |
| Location Privacy | 3,970,000 |
| "Location Privacy" | 29,800 |
| Location Privacy navigation | 1,670,000 |
| "Location Privacy" navigation | 6,330 |
| Location Privacy Threats | 475,000 |
| Location Privacy "Threats" | 7,280 |
| Location Privacy Risks | 2,710,000 |
| Location Privacy "Risks" | 1,390,000 |
| "Location Tracking" and "Privacy" | 25,500 |
| "Location Tracking" and "Privacy Risks" | 1,090 |

| | |
|---|---|
| Location Privacy Protection | 2,730,000 |
| Location Privacy “Protection” | 16,300 |
| “Location Tracking” and “Privacy Protection” | 3,580 |

*Table 1: Table of search term and number of hits*

## 3.6 Practical Screening for Inclusion

Practical screening for inclusion relates to the process of selecting the most relevant papers. Petersen, Feldt, Mujtaba, and Mattsson (2008) identified that the most effective screening method was the use of key word searches that link directly to the research question. Google scholar was found to be highly efficient when screening the search results. It was found that the search result sorting by relevance greatly improved the quality of the results returned. Additionally, Google Scholar displays the number of citations for each research paper returned allowing a quick exclusion of a large set of results that do not meet the citation count threshold.

The search results were analysed using the following article selection criteria:

- Relevant to the five topics outlined in the literature search process section
- Published in a recognised academic journal
- Returned within the first 2 pages of the search results (sorted by relevance)
- Greater than 10 citations
- Paper written in English

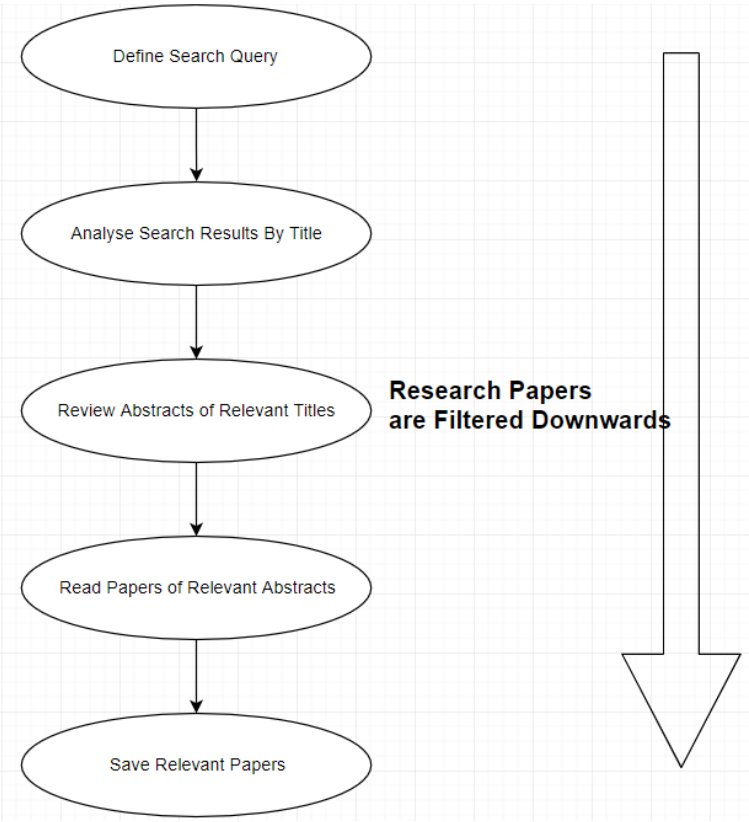


*Figure 23: Paper Selection Flow*

Figure 23 describes the paper selection flow used throughout this study to select the most relevant research. Initially the title of the paper was viewed to gauge the relevance to the topic. If a title was deemed relevant, then the abstract was reviewed to identify more information on the paper. If the abstract was still deemed to be relevant to the topic, the paper was reviewed to identify if it should be saved for further analysis for use in the study. As illustrated above the number of research papers were filtered downwards, with each step removing irrelevant papers. The filtering process used in the research was adopted from Petersen et al. (2008) systemic mapping process for software engineering which is illustrated in Figure 24.

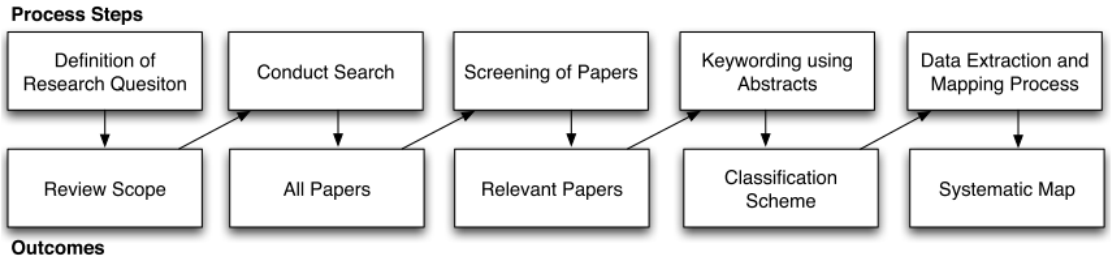


*Figure 24: Systematic Mapping Process (Petersen et al., 2008)*

## 3.7 Quality Appraisal for Exclusion

The practical screening for inclusion step ensures that only relevant literature to the research question are considered. The next stage of the systematic literature review is to consider the quality of the literature, Okoli (2015) identified the quality appraisal stage as having two key purposes:

1. Prioritize papers according to quality
2. Remove papers that are not useful.

One of the key appraisal techniques was the evaluation of the research methodology chosen for each literature. Through reviewing the research methodology employed many papers saved in the previous steps were removed due to their inferior methodological quality. For example, many papers which collected primary data results were removed due to inadequate statistical significance. Keele (2007) identified three principles for quality appraisal based on the research methodology:

1. Bias
   Bias refers to the tendency to produce results that differ from the true result of the study. Study require a systematic methodology to ensure unbiased results.

2. Internal Validity
   The extent to which the research design is suitable. The correct research design minimises the systematic error of the research thus providing validation for the research.

3. External Validity
   The extent to which the research results observed correlate with similar research done in the field. Research papers which have applicable findings to similar studies provide further validation that the research is of a high standard.

## 3.8 Data Extraction

The next step of the systematic literature review process was data extraction, the identification and collection of relevant information from the saved literature. Keele (2007) suggests that the data extraction step must collect all information required to address the research question and answer the research objectives outlined. To answer the three objectives of the research the data extraction was spilt into three distinct sections:

1. Literature related to navigation applications
2. Literature related to location privacy threats
3. Literature related to location privacy protections

The use of EndNote citation software greatly assisted in the data extraction step. Endnote allowed all the citations and PDF files of the literature to be kept together in specific literature groups, making it easy to store and retrieve literature and citation data related to each of the three areas listed above.

## 3.9 Synthesize Studies

The next step of the systematic literature review was to synthesize the saved literature. Okoli (2015) defined the step as the combination of related literature to make 'sense out of their (often large) number'. The literature is aggregated, grouped and compared to critically analyse and combine related information. Keele (2007) identified that the combined output of the literature can be descriptive or narrative in nature. For narrative information it is important that the results from studies are consistent. The result of the synthesis of the studies is reflected in the literature review chapter layout with five main sections outlined:

1. Location Technologies
2. Navigation Devices
3. Multi Modal Transportation Systems
4. Privacy Threats
5. Privacy Protections

## 3.10 Writing the Review

The final step for creating the systematic literature was the writing of the review. The literature review contains all the data and narrative information found, analysed and evaluated through the combination of the previous seven steps of creating a systematic literature review. Kitchenham et al. (2009) argues that if the previous seven steps are completed systematically then the process of writing should be clear with the chapters and main ideas of the literature categorised and understood to a very high standard.

The structure of the write-up was decided based on the objectives for the study. The first three chapters provide extensive background information on how navigation applications and how they impact on location privacy. The chapters on privacy threats and privacy protection outline the main threats posed and the protections in place to protect navigation application users' privacy.

## 3.11 Limitations of Methodology

A Systematic Literature Review depends on the critical analysis of secondary data and research to provide an understanding of a specific area or process, the methodology does not utilise any primary data collection methods (Fink, 2005). Keele (2007) argues that systematic literature reviews can be limited by the quality of the primary data studies of the literature.

One of the core attributes of a systematic literature reviews, is that it can be repeated by other researchers (Keele, 2007), however, Jorgensen and Shepperd (2007) argues that there is no standardized practice for searching for related literature, indicating that there is some level of researcher bias when selecting relevant papers. This research was completed by one researcher, Okoli (2015) and Keele (2007) argue that systematic literature reviews are best suited to research endeavours which have two or more researchers to improve quality and repeatability of the literature review process. Okoli (2015) suggests one researcher proposes the paper with the second critiquing and assessing the validity of the paper, ensuring it meets the quality criteria outlined for the literature review.

## 3.12 Lessons Learnt

The systematic literature review was a very time and energy consuming process. Much time was spent identifying and evaluating the literature before it could be included in the literature review. There were some difficulties in identifying relevant papers, the abstract of many information systems papers was poor requiring further reading to determine their relevance to the research question. Similarly, the literature inclusion and quality appraisal criteria were difficult to correctly define to select the best papers, with too strict rules the literature volume would be extremely limited.

Citation software proved to be a critical component in producing the systematic literature review. Through the use of Endnote, the literature and citations were all managed in a grouped environment. The grouping of literature made it much easier and efficient to analyse and contrast research papers in the same section of the literature review.

# 4 Conclusion

This chapter will discuss the conclusions drawn from the completion of this study. The chapter will illustrate how the research question and objectives have been answered. The chapter concludes by identifying limitations of the research and potential future research in the area of navigation privacy.

## 4.1 Answering the Research Questions

The aim of the research was to provide a holistic examination and analysis of how the use of navigation applications impact upon location privacy. To answer the research question three research objectives were outlined:

1. Gain an understanding of how navigation applications work
2. Identify the potential privacy threats imposed through the use of navigation applications
3. Identify the privacy protections in place to protect users of navigation applications

## 4.2 How Navigation Applications Work

Navigation applications were discussed in detail in the systematic literature review, with three of the five literature subchapters dedicated to explaining enabling technology, devices and the use of multi modal travel planning for navigation applications.

### *4.2.1 Tracking Technologies*

Location tracking technologies were found to be a key enabler of navigation applications, they provide the technological capability to precisely locate a device anywhere on the planet. A number of different tracking methods were identified throughout the literature review documented in Chapter 2.1. The chapter is summarized in Table 2 below which outlines the main advantages and disadvantages found with each location tracking method.

| Tracking Method | Advantages | Disadvantages |
|---|---|---|
| Global Navigation Satellite System (GNSS) | High Precision<br>High Availability<br>High Range<br>Large volume of GNSS receivers<br>Innate Location Privacy, flow is unidirectional | Slow TTFF<br>Obstructed by Environment<br>Performance dependent on Government funding |
| Mobile Network Location Tracking | Fast TTFF<br>Uses existing mobile infrastructure<br>Innate Location Privacy, flow is unidirectional | Low Precision<br>Heavily dependent on Antenna density |
| Wi-Fi Positioning System | Fast TTFF<br>High Precision<br>Indoor Navigation | Location Data shared<br>Low Range<br>Dependent on user mapping |
| Hybrid | Uses best aspects of other tracking methods | Location Data May be shared |

*Table 2: Location Tracking Methods Summary*

### *4.2.2 Navigation Devices*

Navigation devices were identified as any device used to consume navigation applications. Three navigation devices were identified throughout the literature review documented in Chapter 2.2. The chapter is summarized in Table 3 below which outlines the advantages and disadvantages found with each navigation device.

| Device Type | Advantages | Disadvantages |
|---|---|---|
| Portable Navigation Device (PND) | High Privacy – Route calculated on device<br>Cheap | Limited to GNSS Positioning<br>No real time route Information<br>Single purpose device |
| In-Dash Navigation | High Privacy – Route calculated on device<br>Built into vehicle display | Limited to GNSS tracking<br>No real time route Information |
| Smartphones | Hybrid Positioning<br>Real time route Information<br>Multi-faceted Device | Low Privacy – Route calculated by third-party |

*Table 3: Navigation Device Summary*

### *4.2.3 Multi Modal Transportation Systems*

Chapter 2.3 is dedicated to multi modal transportation systems. Multi Modal transportation systems use multiple forms of transportation when generating the optimal route for a user. As outlined in the chapter there are a number of complexities with integrating several transportation systems into a single consolidated, holistic system which is capable of generating travel routes across different forms of transport and transportation operators. Route generation was identified as a key attribute to multi modal navigation applications. Navigation applications are required to generate a number of routes for the user based on two inputs: current location and destination location. Two key facets of route generation were identified: the use of shortest-path algorithms to identify the route between two nodes on a road network and secondly, the impact the form of transportation has on the route defining what transport is applicable for the user. Modern navigation applications generate and adopt the route path in real time by tracking the user's location and using information about the route's conditions to evaluate the best possible route throughout the duration of the journey.

## 4.3 Privacy Threats

The second research objective was to identify privacy threats posed through the use of navigation applications. The privacy threats posed by the use of navigation applications is discussed in detail in Chapter 2.4.

### *4.3.1 Rise of Location Aware Computing*

The rise of location aware computing and location-based services was identified as a major threat to location privacy. Service providers are continually monitoring and capturing sensitive location information which they use to monetise and improve their service offering. The excessive capture of personal data leads to the loss of control for the user.

### *4.3.2 Risk of User Profiling*

User profiling was identified as a significant threat to user privacy. The excessive location data practices enabled by location aware computing allows data collectors to understand traits, tendencies and locations of the user, increasing their risk of user profiling and identification. As highlighted by Barkhuus and Dey (2003) ‘identification’ is the core principle to privacy, if a user can be identified their privacy is compromised.

### *4.3.3 Data Breaches*

Data breaches were identified as a significant threat to user privacy. Data breaches expose sensitive personal information to unauthorized users which results in a loss of control of personal data. The control of personal data is one of the defining principles of privacy, with the loss of control exposing the privacy of individuals (Westin, 1967).

### *4.3.4 Location Privacy Threats*

The capture of location data poses location specific threats to users. Wernke, Skvortsov, Dürr, and Rothermel (2014) identified location specific threats as ‘Knowledge’ based attacks as the perpetrator has access and knowledge of the location patterns of the consumer, compromising their privacy and safety.

### *4.3.5 Privacy Threats Summary*

The privacy threats identified all relate to the data capture and processing of personal information. Through the constant sharing of location information with navigation applications the users privacy is vulnerable through user profiling, data breaches and

location specific attacks which all stem from the emergence of location aware computing and location-based services.

## 4.4 Privacy Protections

The third research objective was to identify privacy protections in place to protect the users of navigation applications. The privacy protections in place to defend users against the privacy threats posed are discussed in detail in Chapter 2.5.

### *4.4.1 Legislation Protections*

Legislation was identified as a key privacy protector. Legislation refers to laws governing how organisations and data controllers capture, process, analyse and use personal information. The leading legislative in the world was identified as the General Data Protection Regulation (GDPR) by the European Union (Buttarelli, 2016). GDPR emphasises that privacy is a fundamental right and protects the user with strict rules and regulations governing how personal data is processed. GDPR enforces harsh penalties on data controllers who fail to comply with the regulation.

### *4.4.2 Privacy Policy*

Privacy policies protect users by informing them of the data management practices of service providers. Privacy policies ensure that the five principles for fair information gathering practices: Notice, Choice, Access, Security and Enforcement are upheld by service providers. As identified in Chapter 2.5.2 concerns have been raised over the nature of privacy policies with many deeming them insufficient for informing and protecting users against data collection practices that may compromise their privacy (Cakebread, 2017) (Wachter, 2018).

### *4.4.3 De-Identification of Personal Data*

As discussed in the privacy threats section the overarching threat to privacy is the capture and storage of personal data. De-identification, anonymisation and encryption are critical privacy protection techniques used to create a pseudonymised dataset. Pseudonymised datasets greatly reduces the risk of user's information being identified, through the authorised or unauthorised disclosure of user data (Esayas, 2015).

### *4.4.4 Mobile Application Permissions*

The majority of navigation applications are consumed on smartphones (Panko, 2018), Smartphones contain powerful sensors which are capable of tracking user movement. Additionally, smartphone stores sensitive personal information which needs to be protected against unauthorised use by downloaded third party applications. Mobile phones provide application permission frameworks for limiting access to sensitive information resources. Users must provide each downloaded application with the consent to track or use information on the device, allowing them to take control of what data mobile applications can access and use.

### *4.4.5 Location Privacy Preserving Mechanisms*

Location privacy preserving mechanisms (LPPM) were identified as location specific protections which obfuscate the location of the user. The goal of LLPM is to reduce the risk of user profiling from Location-Based Quasi-identifiers (Freudiger, Shokri, & Hubaux, 2011). One of the leading LLPM identified in the literature review was the use of position dummies. The technique ensures that service providers cannot be certain which is the real position of the user as the location tracking device sends multiple fake locations to the service provider along with the real location of the user. The literature review also uncovered ‘Mix Zones’ and ‘Cloaking Region’ techniques for reducing location data accuracy (Kasori & Sato, 2015) (Wernke et al., 2014).

### *4.4.6 Privacy Protections Summary*

As identified by Chen et al. (2017) the privacy risk is heightened through the capture and processing of data by third party navigation applications. Global Navigation Satellite System positioning is by default private in nature as the data is unidirectional, the GNSS receiver does not relay back the location of the device to the satellite. The best privacy protection mechanism is to use navigation applications which calculate the route on the device itself and does not use third party route planners. However, as noted by Xu, Teo, Tan, and Agarwal (2009) there is a push pull complex to location service providers, as the privacy increases the quality of service decreases. Built-in route planning systems do not provide any real time information on the route conditions, which may lead to an unpleasant journey experience.

The privacy protections identified in this systematic literature review can be spilt into two categories: Legislative Protections and Technological Protections. The legislative protections enforce certain data practice standards upon data controllers, while,

technological protections refer to the technological techniques and processes used to protect users' data and privacy.

## 4.5 Limitations of Research

As discussed in chapter 3.11 there are a number of limitations to a systematic literature review (SLT) as a research approach:

1. SLT depends solely on the systematic analyse of secondary data and research, there is no primary data collected or analysed.
2. SLT protocol is difficult to implement as a sole researcher.
3. SLT may be broad in nature providing a narrative approach rather than quantitative or qualitive.

## 4.6 Future Research

While navigation applications are not a new concept, the way we consume them through the use of smartphones is relatively new. While there has been a lot of technical research done on the topic of location privacy threats and protections, there has been a lack of research on consumers opinions and knowledge of location privacy. More qualitative analysis needs to be carried out to gain an understanding of consumers knowledge and perception of location privacy.

The legislation protections described in the paper mainly revolve around General Data Protection Regulation. As identified by Newman (2015) GDPR stands in contrast to other privacy protection legislations namely the United States and Canada which have weak privacy protection regulation. More research needs to be done on the impact regulation has on user privacy on a global scale.

## References


Abou-Zeid, M., Witter, R., Bierlaire, M., Kaufmann, V., & Ben-Akiva, M. (2012). Happiness and travel mode switching: findings from a Swiss public transportation experiment. *Transport Policy, 19*(1), 93-104.

Albrecht, J. P. (2016). How the GDPR will change the world. *Eur. Data Prot. L. Rev., 2*, 287.

Alepis, E., Politou, E., & Patsakis, C. (2018). Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions. *Journal of Cybersecurity, 4*(1). Retrieved from https://doi.org/10.1093/cybsec/tyy001. doi:10.1093/cybsec/tyy001

Almuhimedi, H., Schaub, F., Sadeh, N., Adjerid, I., Acquisti, A., Gluck, J., . . . Agarwal, Y. (2015). *Your location has been shared 5,398 times!: A field study on mobile app privacy nudging.* Paper presented at the Proceedings of the 33rd annual ACM conference on human factors in computing systems.

Ambrose, J. K., Bukovsky, D. J., Sedlak, T. J., & Goeden, S. J. (2009, 24-24 April 2009). *Developing a travel route planner accounting for traffic variability.* Paper presented at the 2009 Systems and Information Engineering Design Symposium.

Barkhuus, L., & Dey, A. K. (2003). *Location-Based Services for Mobile Telephony: a Study of Users' Privacy Concerns.* Paper presented at the Interact.

Bates, C. (2015). Basic Antenna Parameters. Retrieved from https://slideplayer.com/slide/8562653/

Beresford, A. R., & Stajano, F. (2003). Location privacy in pervasive computing. *IEEE Pervasive Computing*(1), 46-55.

Beresford, A. R., & Stajano, F. (2004). *Mix zones: User privacy in location-aware services.* Paper presented at the IEEE Annual Conference on Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second.

Besmer, A., Watson, J., & Lipford, H. R. (2010). *The impact of social navigation on privacy policy configuration.* Paper presented at the Proceedings of the Sixth Symposium on Usable Privacy and Security.

Bettini, C., Wang, X. S., & Jajodia, S. (2005). *Protecting privacy against location-based personal identification.* Paper presented at the Workshop on Secure Data Management.

Brandeis, L., & Warren, S. (1890). The right to privacy. *Harvard law review, 4*(5), 193-220.

Buttarelli, G. (2016). The EU GDPR as a clarion call for a new global digital gold standard. *International Data Privacy Law, 6*(2), 77-78. Retrieved from https://doi.org/10.1093/idpl/ipw006. doi:10.1093/idpl/ipw006

Cakebread, C. (2017). You're not alone, no one reads terms of service agreements. Retrieved from https://www.businessinsider.com/deloitte-study-91-percent-agree-terms-of-service-without-reading-2017-11?r=US&IR=T.

Chang, I.-C., Tai, H.-T., Yeh, F.-H., Hsieh, D.-L., & Chang, S.-H. (2013). A VANET-based A* route planning algorithm for travelling time-and energy-efficient GPS navigation app. *International Journal of Distributed Sensor Networks, 9*(7), 794521.

Chen, C. L. P., Zhou, J., & Zhao, W. (2012). A Real-Time Vehicle Navigation Algorithm in Sensor Network Environments. *IEEE Transactions on Intelligent Transportation Systems, 13*(4), 1657-1666. doi:10.1109/TITS.2012.2201478

Chen, L., Thombre, S., Järvinen, K., Lohan, E. S., Alén-Savikko, A., Leppäkoski, H., . . . Kuusniemi, H. (2017). Robustness, Security and Privacy in Location-Based Services for Future IoT: A Survey. *IEEE Access, 5*, 8956-8977. doi:10.1109/ACCESS.2017.2695525

Chitkara, S., Gothoskar, N., Harish, S., Hong, J. I., & Agarwal, Y. (2017). Does this app really need my location?: Context-aware privacy management for smartphones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 1*(3), 42.

Cisco. (2016). *Cisco VNI Global Mobile Data Traffic Forecast, 2016-2021*. Retrieved from

Clarke, R., & Wigan, M. (2011). You are where you've been: the privacy implications of location and tracking technologies. *Journal of Location Based Services, 5*(3-4),

138-155. Retrieved from https://doi.org/10.1080/17489725.2011.637969. doi:10.1080/17489725.2011.637969

Clearinghouse, P. R. (2019). Data Breaches. Retrieved from https://www.privacyrights.org/data-breaches

Commision, D. P. (2018). GDPR - Data Protection Commission - Ireland. Retrieved from https://www.dataprotection.ie/docs/GDPR/1623.html

Commission, P. P. (2017). Act on the Protection of Personal Information. Retrieved from https://www.ppc.go.jp/en/legal/.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*: Sage publications.

Crotts, J., & Zehrer, A. (2012). *An Exploratory Study of Vacation Stress* (Vol. 17).

Dalenius, T. (1986). Finding a needle in a haystack or identifying anonymous census records. *Journal of official statistics, 2*(3), 329.

Damiani, M. L., & Cuijpers, C. (2013, 3-6 June 2013). *Privacy Challenges in Third-Party Location Services.* Paper presented at the 2013 IEEE 14th International Conference on Mobile Data Management.

De Hert, P., Papakonstantinou, V., Malgieri, G., Beslay, L., & Sanchez, I. (2018). The right to data portability in the GDPR: Towards user-centric interoperability of digital services. *Computer Law & Security Review, 34*(2), 193-203. Retrieved from http://www.sciencedirect.com/science/article/pii/S0267364917303333. doi:https://doi.org/10.1016/j.clsr.2017.10.003

Dictionary, O. (2019a). Definition of consent in English. Retrieved from https://en.oxforddictionaries.com/definition/consent

Dictionary, O. (2019b). Definition of privacy in English. Retrieved from https://en.oxforddictionaries.com/definition/privacy.

Dictionary, O. (2019c). Definition of research in English. Retrieved from https://en.oxforddictionaries.com/definition/research

Dürr, F., Skvortsov, P., & Rothermel, K. (2011, 21-25 March 2011). *Position sharing for location privacy in non-trusted systems.* Paper presented at the 2011 IEEE International Conference on Pervasive Computing and Communications (PerCom).

Esayas, S. (2015). The role of anonymisation and pseudonymisation under the EU data privacy rules: beyond the 'all or nothing'approach. *European Journal of Law and Technology, 6*(2).

Essl, T. (2017). Create a Heat Map from your Google Location History in 3 easy Steps. Retrieved from https://towardsdatascience.com/create-a-heat-map-from-your-google-location-history-in-3-easy-steps-e66c93925914

Evangelatos, S., Kalampoukis, Z., Fergadioti, I., Christofi, S., Karakostas, B., & Zorgios, Y. (2017, 3-6 July 2017). *Service availability analysis of a multimodal travel planner using Stochastic Automata.* Paper presented at the 2017 IEEE Symposium on Computers and Communications (ISCC).

Fan, D., & Shi, P. (2010). *Improvement of Dijkstra's algorithm and its application in route planning.* Paper presented at the 2010 seventh international conference on fuzzy systems and knowledge discovery.

Fink, A. (2005). *Conducting Research Literature Reviews: From the Internet to Paper*: SAGE Publications.

Forum, W. E. (2017). Valuing Personal Data and Rebuilding Trust. Retrieved from https://www.weforum.org/whitepapers/valuing-personal-data-and-rebuilding-trust.

Fowler, G. A. (2018). Hands off my data! 15 default privacy settings you should change right now. Retrieved from https://www.seattletimes.com/business/hands-off-my-data-15-default-privacy-settings-you-should-change-right-now/

Freudiger, J., Shokri, R., & Hubaux, J.-P. (2011). *Evaluating the privacy risk of location-based services.* Paper presented at the International conference on financial cryptography and data security.

Geisberger, R. (2011). *Advanced route planning in transportation networks.* Verlag nicht ermittelbar,

GISGeography. (2019). Trilateration vs Triangulation – How GPS Receivers Work. Retrieved from https://gisgeography.com/trilateration-triangulation-gps/

Goddard, M. (2017). The EU General Data Protection Regulation (GDPR): European regulation that has a global impact. *International Journal of Market Research, 59*(6), 703-705.

Google. (2018). Google Developers. Retrieved from https://developers.google.com/android/reference/com/google/android/gms/location/ActivityRecognitionClient

Government, U. (2019). GPS.gov: Space Segment. Retrieved from https://www.gps.gov/systems/gps/space/

Han, D., Jung, S., Lee, M., & Yoon, G. (2014). Building a Practical Wi-Fi-Based Indoor Navigation System. *IEEE Pervasive Computing, 13*(2), 72-79. doi:10.1109/MPRV.2014.24

Harlan, C. (2015). Does MapQuest still exist?' Yes, it does, and it's a profitable business. Retrieved from https://www.washingtonpost.com/business/economy/does-mapquest-still-exist-as-a-matter-of-fact-it-does/2015/05/22/995d2532-fa5d-11e4-a13c-193b1241d51a_story.html?utm_term=.4faefd3bb1c3

Hasan, O., Habegger, B., Brunie, L., Bennani, N., & Damiani, E. (2013, 27 June-2 July 2013). *A Discussion of Privacy Challenges in User Profiling with Big Data Techniques: The EEXCESS Use Case.* Paper presented at the 2013 IEEE International Congress on Big Data.

Heywood, D. I., Cornelius, S. C., & Carver, S. (2011). *An introduction to geographical information systems*: Pearson Prentice Hall.

Hofmann-Wellenhof, B., Lichtenegger, H., & Collins, J. (2012). *Global positioning system: theory and practice*: Springer Science & Business Media.

Huang, K. L., Kanhere, S. S., & Hu, W. (2010). Preserving privacy in participatory sensing systems. *Computer Communications, 33*(11), 1266-1280.

Hughes, W. J. (2017). *Global Positioning System (GPS) Standard Positioning Service (SPS) Performance Analysis Report*. Retrieved from

IBM. (2018). *Examining the 2018 Cost of a Data Breach*. Retrieved from

Jafarnia-Jahromi, A., Broumandan, A., Nielsen, J., & Lachapelle, G. (2012). GPS vulnerability to spoofing threats and a review of antispoofing techniques. *International Journal of Navigation and Observation, 2012*.

Jafri, R., Alkhunji, A. S., Alhader, G. K., Alrabeiah, H. R., Alhammad, N. A., & Alzahrani, S. K. (2013, 11-12 Dec. 2013). *Smart Travel Planner: A mashup of travel-related web services.* Paper presented at the 2013 International Conference on Current Trends in Information Technology (CTIT).

Jorgensen, M., & Shepperd, M. (2007). A systematic review of software development cost estimation studies. *IEEE Transactions on software engineering, 33*(1), 33-53.

Kapadia, A., Triandopoulos, N., Cornelius, C., Peebles, D., & Kotz, D. (2008). *AnonySense: Opportunistic and privacy-preserving context collection.* Paper presented at the International Conference on Pervasive Computing.

Karagiorgou, S., & Pfoser, D. (2012). *On vehicle tracking data-based road network generation.* Paper presented at the Proceedings of the 20th International Conference on Advances in Geographic Information Systems.

Kasori, K., & Sato, F. (2015, 2-4 Sept. 2015). *Location Privacy Protection Considering the Location Safety.* Paper presented at the 2015 18th International Conference on Network-Based Information Systems.

Keele, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*. Retrieved from

Kido, H., Yanagisawa, Y., & Satoh, T. (2005). *An anonymous communication technique using dummies for location-based services.* Paper presented at the ICPS'05. Proceedings. International Conference on Pervasive Services, 2005.

Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology, 51*(1), 7-15. Retrieved

from http://www.sciencedirect.com/science/article/pii/S0950584908001390. doi:https://doi.org/10.1016/j.infsof.2008.09.009

Kooistra, J. (2018). Newzoo's 2018 Global Mobile Market Report: Insights into the World's 3 Billion Smartphone Users. Retrieved from https://newzoo.com/insights/articles/newzoos-2018-global-mobile-market-report-insights-into-the-worlds-3-billion-smartphone-users/

Krumm, J. (2009). A survey of computational location privacy. *Personal and ubiquitous computing, 13*(6), 391-399.

Leontiadis, I., Efstratiou, C., Picone, M., & Mascolo, C. (2012). *Don't kill my ads!: balancing privacy in an ad-supported mobile application market.* Paper presented at the Proceedings of the Twelfth Workshop on Mobile Computing Systems & Applications.

Lin, J., Amini, S., Hong, J. I., Sadeh, N., Lindqvist, J., & Zhang, J. (2012). *Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing.* Paper presented at the Proceedings of the 2012 ACM conference on ubiquitous computing.

Lin, J., Liu, B., Sadeh, N., & Hong, J. I. (2014). *Modeling users' mobile app privacy preferences: Restoring usability in a sea of permission settings.* Paper presented at the 10th Symposium On Usable Privacy and Security ({SOUPS} 2014).

Mäenpää, H., Lobov, A., & Martinez Lastra, J. L. (2017). Travel mode estimation for multi-modal journey planner. *Transportation Research Part C: Emerging Technologies, 82*, 273-289. doi:10.1016/j.trc.2017.06.021

Manalo, A. (2019). Best Navigation Apps: Google Maps vs. Apple Maps vs. Waze vs. MapQuest. Retrieved from https://smartphones.gadgethacks.com/how-to/best-navigation-apps-google-maps-vs-apple-maps-vs-waze-vs-mapquest-0194591/

Mateosian, R. (2013). Ethics of Big Data. *IEEE Micro, 33*(2), 60-61. doi:10.1109/MM.2013.35

Molin, E., Mokhtarian, P., & Kroesen, M. (2016). Multimodal travel groups and attitudes: A latent class cluster analysis of Dutch travelers. *Transportation Research Part A: Policy and Practice, 83*, 14-29. Retrieved from http://www.sciencedirect.com/science/article/pii/S0965856415002591. doi:https://doi.org/10.1016/j.tra.2015.11.001

Nations, U. (1948). Universal Declaration of Human Rights. Retrieved from https://www.un.org/en/universal-declaration-human-rights/.

Nations, U. (2018). Personal Data Protection and Privacy Principles. Retrieved from https://www.unsceb.org/CEBPublicFiles/UN-Principles-on-Personal-Data-Protection-Privacy-2018.pdf

Newman, A. (2015). *What the "right to be forgotten" means for privacy in a digital age* (Vol. 347).

Nikolaenko, V., Ioannidis, S., Weinsberg, U., Joye, M., Taft, N., & Boneh, D. (2013). *Privacy-preserving matrix factorization.* Paper presented at the Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security.

Novet, J. (2015). Hands on with Google Maps' Your Timeline: Fascinating, but freaky. Retrieved from https://venturebeat.com/2015/07/27/hands-on-with-google-maps-your-timeline-fascinating-but-freaky/

Okoli, C. (2015). A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems, 37*.

Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research.

P. Nath, N., Parija, S., Sahu, P., & Singh, S. (2015). *Survey Paper: Location Management in CDMA Network* (Vol. 8).

Panko, R. (2018). The Popularity of Google Maps: Trends in Navigation Apps in 2018 *The State of Tech.* Retrieved from https://themanifest.com/app-development/popularity-google-maps-trends-navigation-apps-2018

Paonni, M., Anghileri, M., Wallner, S., Avila-Rodriguez, J.-A., & Eissfeller, B. (2010). Performance assessment of GNSS signals in terms of time to first fix for cold, warm and hot start. *Proc. of the ION ITM*, 25-27.

Papageorgiou, A., Strigkos, M., Politou, E., Alepis, E., Solanas, A., & Patsakis, C. (2018). Security and Privacy Analysis of Mobile Health Applications: The Alarming State of Practice. *IEEE Access, 6*, 9390-9403. doi:10.1109/ACCESS.2018.2799522

(2017). *The Future of Everything* [

PersistenceMarketResearch. (2017). Global Market Study on In-Dash Navigation System: Europe to Remain Dominant in the Market During 2017-2022. Retrieved from https://www.persistencemarketresearch.com/market-research/in-dash-navigation-system-market.asp

Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). *Systematic mapping studies in software engineering.* Paper presented at the Ease.

Ratsameethammawong, P., & Kasemsan, K. (2010). Mobile phone location tracking by the combination of gps, wi-fi and cell location technology. *Communications of the IBIMA*.

Regulation, G. D. P. (2018). Principles relating to processing of personal data. Retrieved from https://gdpr-info.eu/art-5-gdpr/

Rita El, K. (2019). Google Maps hits 5 billion downloads on the Play Store, does it after YouTube but before the Google app. Retrieved from https://www.androidpolice.com/2019/03/09/google-maps-hits-5-billion-downloads-on-the-play-store-does-it-after-youtube-but-before-the-google-app/

Saunders, M., Lewis, P., & Thornhill, A. (2007). Research methods. *Business Students 4th edition Pearson Education Limited, England*.

Schmitt, L., Currie, G., & Delbosc, A. (2014). Lost in transit? Unfamiliar public transport travel explored using a journey planner web survey. *Transportation, 42*(1), 101-122. doi:10.1007/s11116-014-9529-2

Schwab, P.-N. (2018). Reading privacy policies of the 20 most-used mobile apps takes 6h40. Retrieved from https://www.intotheminds.com/blog/en/reading-privacy-policies-of-the-20-most-used-mobile-apps-takes-6h40/

Services, U. S. D. o. H. a. H. (2015). Information Memorandum Retrieved from https://www.acf.hhs.gov/cb/resource/im1504.

Shokri, R., Theodorakopoulos, G., Troncoso, C., Hubaux, J.-P., & Le Boudec, J.-Y. (2012). *Protecting location privacy: optimal strategy against localization attacks.* Paper presented at the Proceedings of the 2012 ACM conference on Computer and communications security.

Simão, J. P. (2015). *Impacts of Advanced Travel Information Systems on Travel Behaviour: Smartmoov'case study.* Paper presented at the ECTRI Young Researchers' Seminar.

Spiekermann-Hoff, S. (2012). The challenges of privacy by design. *Communications of the ACM (CACM), 55*(7), 34-37.

Stevens, G. M. (2012). Data security breach notification laws. In: Congressional Research Service Washington, DC.

Su, X., Tong, H., & Ji, P. (2014). Activity recognition with smartphone sensors. *Tsinghua Science and Technology, 19*(3), 235-249. doi:10.1109/TST.2014.6838194

Tankard, C. (2016). What the GDPR means for businesses. *Network Security, 2016*(6), 5-8.

Tyagi, A. K., & Sreenath, N. (2015, 2-4 April 2015). *Location privacy preserving techniques for location based services over road networks.* Paper presented at the 2015 International Conference on Communications and Signal Processing (ICCSP).

UnitedNations. (1980, 24th May). Paper presented at the United Nations Convention on International Multimodal Transport of Goods, Geneva.

Vicente, C. R., Freni, D., Bettini, C., & Jensen, C. S. (2011). Location-Related Privacy in Geo-Social Networks. *IEEE Internet Computing, 15*(3), 20-27. doi:10.1109/MIC.2011.29

Wachter, S. (2018). Normative challenges of identification in the Internet of Things: Privacy, profiling, discrimination, and the GDPR. *Computer Law & Security Review, 34*(3), 436-449. Retrieved from

http://www.sciencedirect.com/science/article/pii/S0267364917303904. doi:https://doi.org/10.1016/j.clsr.2018.02.002

Wan, B., Wan Bejuri, W. M. Y., Mohamad, M., Mohd, M., & Sapri, M. (2011). *Ubiquitous Positioning: A Taxonomy for Location Determination on Mobile Navigation System* (Vol. 2).

Wernke, M., Skvortsov, P., Dürr, F., & Rothermel, K. (2014). A classification of location privacy attacks and approaches. *Personal and ubiquitous computing, 18*(1), 163-175.

Westin, A. F. (1967). Privacy and freedom.

Wheatley, S., Maillart, T., & Sornette, D. (2016). The extreme risk of personal data breaches and the erosion of privacy. *The European Physical Journal B, 89*(1), 7.

Wilkinson, S. (2018). Brazil's new General Data Protection Law. *Journal of Data Protection & Privacy, 2*(2), 107-115.

Wu, K.-W., Huang, S. Y., Yen, D. C., & Popova, I. (2012). The effect of online privacy policy on consumer privacy concern and trust. *Computers in Human Behavior, 28*(3), 889-897. Retrieved from http://www.sciencedirect.com/science/article/pii/S0747563211002767. doi:https://doi.org/10.1016/j.chb.2011.12.008

Xi, Y., Schwiebert, L., & Shi, W. (2014). Privacy preserving shortest path routing with an application to navigation. *Pervasive and Mobile Computing, 13*, 142-149. Retrieved from http://www.sciencedirect.com/science/article/pii/S1574119213000795. doi:https://doi.org/10.1016/j.pmcj.2013.06.002

Xu, H., Teo, H.-H., Tan, B. C., & Agarwal, R. (2009). The role of push-pull technology in privacy calculus: the case of location-based services. *Journal of management information systems, 26*(3), 135-174.

Yang, J., Varshavsky, A., Liu, H., Chen, Y., & Gruteser, M. (2010). *Accuracy Characterization of Cell Tower Localization*.

Zeinalipour-Yazti, D., Laoudias, C., Georgiou, K., & Chatzimilioudis, G. (2018). Internet-based Indoor Navigation Services. *IEEE Internet Computing*, 1-1. doi:10.1109/MIC.2017.265101954

Zhang, X., Yang, L. T., Liu, C., & Chen, J. (2014). A Scalable Two-Phase Top-Down Specialization Approach for Data Anonymization Using MapReduce on Cloud. *IEEE Transactions on Parallel and Distributed Systems, 25*(2), 363-373. doi:10.1109/TPDS.2013.48

Zhou, Y., Jun, L., & Lamont, L. (2012, 3-7 Dec. 2012). *Multilateration localization in the presence of anchor location uncertainties.* Paper presented at the 2012 IEEE Global Communications Conference (GLOBECOM).

# 5 Appendices

## 5.1 Appendix 1: Navigation Application Comparison (Manalo, 2019)

| Best Maps & Navigation Apps for Mobile | | | | |
|---|---|---|---|---|
| GADGET HACKS | Google Maps | Apple Maps | Waze | MapQuest |
| General | | | | |
| Platform | Android, iOS, macOS, Windows | iOS, macOS | Android, iOS, Windows | Android, iOS, macOS, Windows |
| Map Features | | | | |
| Countries & Terriories Mapped | 266 | 181 | 72 | 252 |
| Countries & Territories with Driving Directions | 256 | 101 | 72 | 252 |
| Street View | Yes | No | No | No |
| Overlays | Satellite, Terrain, Transit, Traffic, Bicycling | Satellite, Transit | None | Satellite |
| 3D View | 3D Structures | 3D Renderings | No | No |
| Live Location Sharing | Yes | Yes | No | No |
| Location History | Yes | Yes | Yes | Yes |
| Cultural Hotspot Indicators | Yes | No | No | No |
| Weather Data | None | Weather, Temperature, Air Quality | None | Weather, Temperature |
| Indoor Maps | Airports, Malls, Museums | Airports, Malls | No | No |
| Offline Maps | Yes | Yes | No | No |

| Navigation Features | | | | |
|---|---|---|---|---|
| Traffic Data | Alternate Routes, Accidents, Road Work, Speed Traps | Alternate Routes, Accidents, Road Work | Accidents, Alt. Routes, Road Work, Potholes, Police, Speed Traps | Accidents, Road Work, Traffic Cameras |
| Traffic Data Source | In-House, User Curated | In-House, Third-Party | In-House, User Curated | Third-Party, User Curated |
| High Traffic Warnings | Yes | Yes | No | No |
| Speed Limits | Yes | Yes | Yes | Yes |
| Lane Guidance | Yes | Yes | Yes | Yes |
| Add Toll & HOV Passes | No | No | Yes | No |
| Avoid Tolls & Highways | Yes | Yes | Yes | Yes |
| Choose Different Routes | Yes | Yes | Yes | Yes |
| Add Pit Stops | Unlimited | 1 | 1 | Unlimited |
| Show Gas Prices | Yes | Yes | Yes | Yes |
| Hands-Free Control in App | Yes | Yes | Yes | No |
| Directions Using Other Modes of Transport | Transit, Biking, Walking, Ride Share | Transit, Walking, Ride Share | Motorcycles, Taxis | Biking, Walking |
| Re-Center | Yes | Yes | Yes | Yes |

| | | | | |
|---|---|---|---|---|
| Accessible Navigation | Yes | Yes | No | No |
| Save Parking Spot | Yes | Yes | No | No |
| Offline Navigation | Yes | Yes | No | No |
| Works With Screen Off | Yes | Yes | Yes | Yes |
| App Features | | | | |
| Dark Mode | Yes | Yes | Yes | Yes |
| Ride Share Integration | Uber, Lyft, Lime, | Uber, Lyft | None | None |
| Picture In Picture | Yes (Android Only) | No | No | No |
| Lock Screen Navigation | Yes | Yes | Yes | Yes |
| Show Festivals & Protests | No | No | Yes | No |
| Personalized Recommendations | Yes | No | No | No |
| Book Dinner Reservations | Via OpenTable | Via OpenTable | No | No |
| Report Traffic Issues | No | No | Yes | No |
| Post Reviews | Yes | No | No | No |
| Car Support | Android Auto, CarPlay | CarPlay | Android Auto, CarPlay | No |

| | | | | |
|---|---|---|---|---|
| AR Features | Interactive Street View | Flyover | None | None |
| Widgets | Yes | Yes | Yes | No |
| Music Integration | Yes | Yes | Yes | No |
| Siri Shortcuts | Yes | Yes | Yes | No |

Navigation Application Comparison (Manalo, 2019)

## 5.2 Appendix 2: Smartphone 3rd party libraries (Lin, Liu, Sadeh & Hong, 2014)

| Type | Examples | Description |
|---|---|---|
| *Utility* | Xmlparser, hamcrest | Utility java libraries, such as parser, sql connectors, etc |
| *Targeted Ads* | admob, adwhirl, | Provided by mobile behavioral ads company to display in-app advertisements |
| *Customized UI Components* | Easymock, kankan, | Customized Android UI components that can be inserted into apps. |
| *Content Host* | Youtube, Flickr | Provided by content providers to deliver relevant image, video or audio content to mobile devices. |
| *Game Engine* | Badlogic, cocos2dx | Game engines which provide software framework for developing mobile games. |
| *SNS* | Facebook, twitter, | SDKs/ APIs to enable sharing app related content on SNSs. |
| *Mobile Analytics* | Flurry, localytics | Provided by analytics company to collect market analysis data for developers. |
| *Secondary Market* | Gfan, ximad, getjar… | Libraries provided by other unofficial Android market to attract users. |
| *Payment* | Fortumo, paypal, zong… | e-payment libraries |

## 5.3 Appendix 3: Systematic Literature Review (Okoli, 2015)