TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Roberta Murniece 201672IVCM

CAN WE TRUST THE GOOGLE MAPS TIMELINE

Master's thesis

Supervisor: Dr. Matthew James Sorell

Tallinn 2022

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Roberta Murniece 201672IVCM

KAS VÕIME USALDADA GOOGLE MAPSI AJATELGE

Magistritöö

Juhendaja: Dr. Matthew James Sorell

Tallinn 2022

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis and this thesis has not been presented for examination or submitted for defence anywhere else. All used materials, references to the literature and work of others have been cited.

Author: Roberta Murniece

11.05.2022

Abstract

Mobile devices are dominant evidence sources in digital forensic cases due to their portability and personal sensitivity. Google Maps is a location-based application pre-loaded on most Android devices and widely used by iOS device owners. After enabling location history, this application stores location data and presents a logical timeline of visits and routes. How can we assure that the presented timeline is actually true? With this preliminary study, we analyzed the accuracy of the data collected and presented in the application. We conducted controlled physical experiments with several devices, adjusted navigation sensors, and walked the same route to determine the accuracy of collected raw location data. We also visited 20 different locations to observe the precision of identifying visited places in the timeline. The research aims to support the forensic investigators with the best practice timeline acquisition protocol. The proposed data acquisition protocol captures and considers the accuracy and limitations of each acquisition layer as observed in the experiments.

Keywords: *location-based applications, Google Maps timeline, mobile device forensics, digital evidence, location, admissibility.*

The thesis is in English language and contains 98 pages of text, 6 chapters, 38 figures, and 28 tables.

Annotatsioon

Mobiilseadmed on digitaalsetes kohtuekspertiisi juhtumites tänu kaasaskantavusele ja tundlikele isikuandmetele olulised tõendite allikad. Google Maps on asukohapõhine rakendus, mis on eellaaditud enamikesse Androidi operatsioonisüsteemiga seadmetesse ja mida kasutavad paljud iOS-operatsioonisüsteemiga seadmete omanikud. Pärast asukoha kronoloogilise loendi lubamist salvestab rakendus asukohaandmed ja esitab külastuste ja marsruutide loogilise ajatelje. Kuidas saame tagada, et esitatud ajatelg on tegelikult tõene? Käesolevas eeluuringus analüüsisime rakenduses kogutud ja esitatud andmete täpsust. Viisime ellu mitme seadmega kontrollitud füüsilised katsed, reguleerisime navigatsiooniandureid ja läbisime sama marsruudi, et määrata asukoha kohta kogutud algandmete õigsust. Samuti külastasime 20 erinevat kohta, et jälgida ajateljel külastatud paikade tuvastamise täpsust. Uuringu eesmärk on toetada kohtuekspertiisi uurijaid parimate tavade kohase ajatelje omandamise protokolliga. Väljapakutud andmete omandamise protokoll hõlmab ja arvestab iga omandatud andmekihi täpsust ja piiranguid, nagu katsetes on täheldatud.

Märksõnad: *asukohapõhised rakendused, Google Mapsi ajatelg, mobiilseadmete kohtuekspertiis, digitaalsed tõendid, asukoht, vastuvõetavus.*

Lõputöö on kirjutatud inglise keeles ning sisaldab teksti 98 leheküljel, 6 peatükki, 38 joonist, 28 tabelit.

Acknowledgments

I want to thank Professor Dr. Matthew James Sorell for supporting me throughout the writing process. Thank you for giving me the idea about the topic. From just an idea, we managed to conduct structured experiments and gather insight about an application that follows every step of our lives. I hope that our paths will cross in the future, and we may research something even more enjoyable. I wish you all the best!

Thank you, Aigars, for being with me from the very beginning. We grow together.

I want to dedicate this research to the women of my family. To my mother Erika, and grandmother Zenta. I am part of you. I would have never done this without cherishing where I come from.

List of abbreviations and terms

iOS	iPhone Operating System
LBS	Location Based Systems
RSS	Received Signal Strength
RTOF	Received Time of Flight
RSSI	Received Signal Strength Indicator
TDOA	Time Difference of Arrival
TOA	Time of Arrival
AOA	Angle of Arrival
KML	Keyhole Markup Language
JSON	JavaScript Object Notation
GPS	Global Positioning System
GIS	Geographic Information System
GSM	Global System of Mobile Communication
GNSS	Global Navigation Satellite System
PPS	Precise Positioning System
SPS	Standard Positioning System
WLAN	Wireless Local Area Network

Table of contents

List of figures

List of tables

1 Introduction

Motivation

Mobile devices are dominant evidence sources in digital forensics due to their portability and personal sensitivity. Google Maps is a location-based application pre-loaded on most Android devices and widely used by iOS device owners. Google Maps application is the leading mapping application [1]. After enabling location history, this application stores location data and presents a logical timeline of visits and routes. Since individuals tend to carry their mobile devices with them most of the time, the location data collected and stored by the Google Maps application is valuable for forensics investigators. However, for data to be valid in court as digital evidence, it must hold a particular criterion of forensic soundness. How can we assure that the presented timeline in the Google Maps is actually true? A 2017 article in the "Times Union" newspaper states that google location evidence in a criminal trial retrieved from the defendant's mobile phone was not suitable for prosecution, with the following statement: "*failed to meet their burden of demonstrating that the science underlying Google location services has gained general acceptance in the relevant scientific community*" [2]. We can see the importance of the data collected and stored in these location-based applications, however, without scientific and peer-reviewed research, this data can be of no value in the court. The primary motivation of the research is to present an insight into the data accuracy collected by the Google Maps application to the forensic community.

We were inspired by the research done at the Netherlands Forensic Institute [3], which evaluated the accuracy of the Google Maps timeline data. Our objective is to reproduce similar experiments of manipulating navigations sensors on the mobile device and extend the scope by including additional research questions.

Novelty

Several methods have been used to test the Google Maps timeline and navigation accuracy [4], [5]. Explicit research has been done by A. Macarulla Rodriguez, C. Tiberius,

R. van Bree, and Z. Geradts [3] at the Netherlands Forensic Institute on the timeline accuracy assessment in terms of navigation settings on the phone. Similarly, we will also analyze Google Maps timeline accuracy based on mobile device navigation sensors. We will extend this research by analyzing visit detection accuracy from the raw data.

The research gap addressed is limited academic literature about the forensically verified data extraction processes from location-based mapping applications. Majority of published mobile device forensics studies focus on location data stored locally on the device [6, pp. 77–78], [7], [8], [9] and less on the location data stored in the cloud accounts. Our study focuses on the data acquisition that is stored in the cloud Google account. We further extend academic research on mobile device cloud application acquisition by explicitly proposing a data acquisition process from the Google Maps timeline.

Research Questions

We aim to answer two research questions to support our proposed the data acquisition process:

I. How often and accurately are raw data points collected by the Google Maps application with different mobile device navigation sensors enabled?

II. What location types are more likely to be detected and presented as visits in the Google Maps timeline?

Research goal

We conduct controlled physical experiments to understand the navigation sensor impact on the Google Maps timeline accuracy as well analyze what locations are more likely to be detected by the mobile application. Based on this insight, we will also recommend a data acquisition process from the Google Maps timeline application on a mobile device. Each acquisition layer will capture and consider data interrogation, use cases, accuracy, and limitations. The proposed process will support academia and digital forensics investigators to quantify and qualify the Google Maps timeline data when presented as evidence in court. Secondly, we will present the best practice timeline acquisition protocol under manual and logical acquisition techniques. The acquisition protocol will benefit forensic investigators in conducting digital investigations. This research can be

used as the first step to conduct more detailed analysis of other widespread location-based applications.

Research Methods

We will use empirical research methodology and observe the data and performance of the Google Maps application from different points of view. Based on the observed phenomena, we will develop a protocol for the best practice data acquisition. We will conduct Google Maps timeline data analysis and ad-hoc experimentation to refine the controlled physical experiments. Quantitative and qualitative timeline data analysis from two user accounts is necessary to understand the points of interest. Ad-hoc experimentation is required to understand the application performance and our ability to influence the outcome. The final set of controlled experiments will be conducted based on the observations from previous experiments and collected data. We will perform two different controlled physical experiments to test the accuracy of the Google Maps timeline in different settings. The empirical research methodology was selected because we will directly observe the performance of the Google Maps timeline that will shape our proposed final data acquisition protocol.

Test data: raw and semantic location history exports from Google accounts on three mobile test devices.

Validation data: First, validation data for navigation sensor impact experiments will be collected from the "GPS Logger Lite" application. The application will be run on Samsung Galaxy S21 mobile device, and it will be used as the ground truth data. The application collects the device's physical location for every second of the experiment. Second, validation data for visit detection accuracy experiments will be manually specified by recording the visited location during the experiment. Physical, address, and semantic location data will be collected.

Raw data exports will be compared with "GPS Logger Lite" detected locations. We will calculate data capture interval, Google proposed accuracy and haversine distance from test to validation data. Semantic location history data will be compared with manually specified visited locations. The data analysis will be performed manually based on the following logic: if the most exact location is presented in the timeline, we apply 1; if a location is near the actual visited location, we apply 0.5; if the visit is not present, we

apply 0. The semantic location detection accuracy result is presented as a success rate in percentage.

To validate the research outcome, three test devices are used in all experiments to compare and contrast the behaviors. Two devices have an Android operating system, and one device has iOS operating system. Average measures from all devices are calculated to present summarized results. All experiment organization steps will be outlined in Chapter 2 to maintain reproducibility.

Scope

Our research will analyze the validity of only one location-based application – Google Maps. This application was chosen because it is the leading mapping application in the United States [1] and it is also extensively used in Europe. Because of the high penetration, our research will benefit most cases where smartphone location data needs to be retrieved from a mobile device. We will only focus on location data extracted from the application and not stored locally on a mobile device. To conduct the research, we will use iOS and Android mobile phones: Samsung Galaxy S21, Samsung Galaxy A7, Huawei P20 Light, and iPhone 7Plus. Both: manual and logical acquisition techniques are used in the proposed data acquisition protocol because solely logical acquisition techniques may be inadmissible in some cases.

The controlled physical experiments will be performed in Latvia, Europe.

Limitations

There are several limitations that we outline for the research:

I. We only use Google Maps timeline. Other mapping applications may provide different results.

II. The research is conducted in specific regions and may not apply globally.

III. Location accuracy depends on a device type. We use a small sample size of 3 test devices. Higher device variation with older and newer models may be needed to extend the research.

IV. Research [3] indicates that location accuracy is influenced by the means of transport, weather, and traffic. These variables are out of the scope of the current research.

V. Google services updates are uncontrolled variables that may affect the research outcome.

VI. We consider the Google Maps timeline in a passive mode and have not analyzed the timeline performance during a journey with enabled active navigation.

Key assumptions

Location spoofing: we assume that there is no deliberate location spoofing, such as GNSS signal hijacking or software intervention.

Mobile phone device condition: we assume that all devices are fully operational and there are no defects which may affect the performance of the Google Maps application.

GNSS and cellular network: we assume that there are no significant outages or other global events that affect the functionality of GNSS or cellular network signals.

Location of the experiments: the location is selected randomly to reflect valid results.

Google account integrity: accounts are assumed to be created only for experiment purposes, and their settings are the same.

Time of the day: we assume that the time of the day does not affect the experiment results because the experiments are conducted at different time intervals.

Ethical considerations

Our research analyzes the personal location data of the author (Murniece) and supervisor (Sorell). The data is transformed and analyzed only in an aggregated form without any personal information disclosure and was used during the initial development phase of the research. Full informed consent of the respective data owners was received. We want to highlight that the Google Maps application is collecting sensitive data, and this data acquisition is limited to everyone except the owner of the Google account.

All concerns and faults in the Google Maps application that we may uncover during the experiments will be communicated with the application owner Google Inc.

Thesis outline

Chapter 1 introduces the research, including motivation, research questions, novelty, and research methods. This chapter presents the importance and the final goal of the research and informs about the key assumptions, limitations, and ethical considerations.

Chapter 2 includes a literature review about Google maps, mobile cloud and location data forensics, and defining forensically sound evidence.

Chapter 3 gives an insight into the research background. We explain the methodology of two controlled experiments, give insight into mobile device navigation techniques tested in the experiments, and explain geographic information systems. We also present a brief analysis and structure of the raw and semantic location history data retrieved from the Google Maps application.

Chapter 4 presents the results of the performed experiments.

Chapter 5 explains our proposed data acquisition protocol.

Chapter 6 includes the final conclusions and suggested future research.

2 Literature review

2.1 Google Maps and reverse geocoding

Netherlands Forensic Institute [3] has reported extensive research on the Google Maps timeline accuracy assessment and error prediction. The research analyses how different external variables and mobile device configuration affect the Google Maps timeline data accuracy compared to a superior location retrieving device. The results show that GPS configuration holds the highest accuracy, following 3G, 2G, and the least accurate is when only Wi-Fi is enabled. Additionally, the research reports a linear regression model that predicts the accuracy level based on the actual location and environmental variables.

A comparison of accuracy between Google Maps and OpenStreetMaps was presented by Cipeluch et al. [4]. The technique used was to compare points of interest in Ireland and overlay KML files from Google Maps and Bing Maps on OpenStreetMap. The accuracy of each application was evaluated based on completeness, the correctness of the map compared to superior knowledge, and spatial information value.

Google Maps platform defines reverse geocoding as "the process of converting geographic coordinates into a human-readable address" [10]. Mapbox geocoding API states that forward geocoding converts location text to a geographic coordinate system, and reverse geocoding performs the opposite [11]. Figure 1 presents three location data types: physical location, logical location, and semantic location.

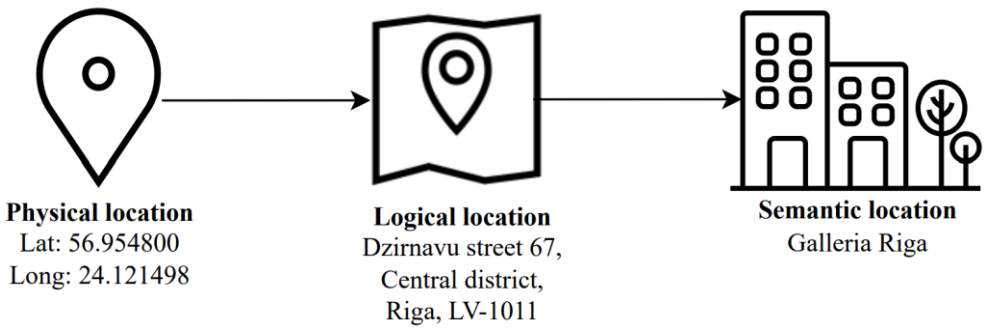
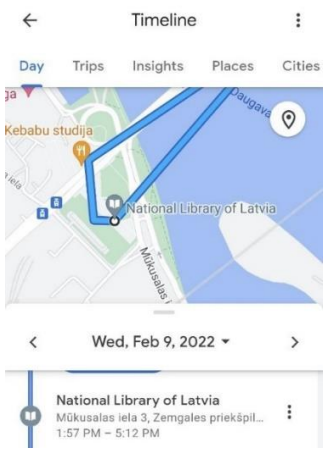
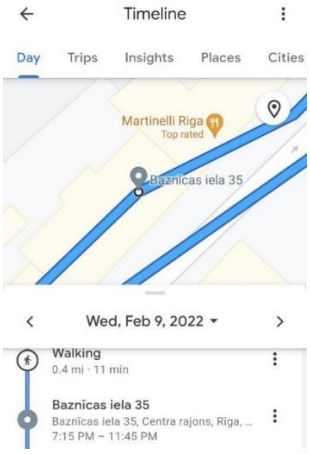


Figure 1. Location data types and examples.

Physical location is the most precise location type, and it depicts location as latitude and longitude values. Increasing the decimal degrees of the values increases the precision. Logical and semantic locations are associated with addresses, buildings, and physical landmarks. Several physical data points can be linked with one logical or semantic location. As mentioned by F. Gülgen and B. Kiliç (2020) [5], geocoding services' significant problem is linking unique physical location points to addresses on the geospatial database. Not all logical locations have an associated semantic location. The Google Maps timeline primarily presents a visit as a semantic location; however, if it is unavailable, then the logical location of the visit is presented.

Hightower and Borriello (2001) [12] splits the position of the device into two subgroups: physical position and symbolic location. The symbolic location is an abstract idea of the location, such as the mall, and the physical location is the object's actual position, such as latitude and longitude. The symbolic location information can be derived from the physical location data, thereby creating a high dependency on the resolution of the physical positioning system. They mention that the system that can provide physical location can also be augmented to provide symbolic location information and infrastructure. Two location types in Google Maps timeline are presented in Table 1.

Table 1. Logical and semantic location in Google Maps timeline.

Semantic location **"National Library of Latvia"**	**Logical location** **"Baznicas street 35, Riga, Latvia."**
	

Reverse geocoding is an integral part of location-based services. The Google Maps timeline accuracy relies on the accuracy of the used forward and reverse geocoding methods. There are various online geocoding service application programming interfaces

(API) available, for example, Google Maps, Nominatim, ArcGIS, Yahoo! Maps, Bing Maps, Geonames, and Area Mean [13]. [14] in 2011 presents the evaluation of the accuracy of Google Maps geocoding services in a large city of Belo Horizonte, Brazil, and it shows that Google Geocoding API found 99.9% of input addresses. Additionally, 74.6% of all addresses were geocoded on an address level, and 73% of all returned locations were within 150 meters from the input position. Gülgen and Kiliç [5] in 2020 compared the accuracy of Google Maps and Bing Maps geocoding services in Miami Beach and Fatih regions. The positional accuracy was measured as a straight line distance in the Euclidean space between the reference point and the point obtained through online geocoding in mapping applications. The results present that the mean of positional accuracy for Google Maps is 12.2 meters in Miami Beach and 8.3 meters in Fatih, for Bing Maps 12.9 and 13.9 respectively [5]. Eight different free online geocoding services were compared by O. Kounadi, T. J. Lampoltshammer, M. Leitner, and T. Heistracher in Vienna, Austria in 2013 [13]. The precision measurement was based on the percentage of the input and output locations that were the same or almost identical.

Current consumer technology utilizing geocoding is rapidly changing, hence some academic research performed more than five years ago may not reflect the current state. Due to these technological advancements the accuracy of the geocoding capabilities may be improving.

2.1.1 Semantically meaningful location mining

Standard mobile devices are integrated with GNSS receivers and generate substantial raw geographic coordinate data streams. The raw data received by the device and applications do not hold any meaning for an average user. Relevant algorithms and interpretations must be applied to extract visited or essential locations from the raw data streams and present them to the used in the Google Maps application. Tobler's first law of geography, "everything is related to everything else, but near things are more related than distant things" [15], can be used as an assumption that individuals with similar location history patterns may hold other related behavioral patterns. [16] presents a hierarchical-graph-based approach to calculate the similarity of the individual users based on their past location data. The research also implies that similarity can also present the correlation between geographical locations.

Various techniques for mining semantically meaningful locations to the user from raw data streams have been generally researched in [17]–[21]. The main goal of these papers is to find the most applicable techniques to retrieve the most precise semantic locations. A predictive model based on users' past movements in raw location data format was developed to present semantically meaningful locations to the user. The model learns from the past data and predicts the users' possible movements with a 48.4-meter distance to the center accuracy. The research was performed in 2003, and it states that "such predictive models might become an integral part of intelligent wearable agents" [18].

X. Cao, G. Cong, and C. S. Jensen in [19] present methods on extracting significant semantically meaningful locations from vast amounts of raw location data. The indicators of significance used in the research are the number and the duration of the visits and the distance traveled. An essential aspect of extracting semantically meaningful data is the level of significance of each data point to the selected context, for example, a particular user group or period. The research presents these semantic locations as "Top-k hot semantic locations". Other factors proposed in the framework are different relationships and distances between the semantic locations.

The research [21] presents three techniques of place discovery from raw location data: time-based clustering technique (a place is significant when the distance from the new location is above threshold and time spent in the new location is above the threshold), partition clustering (K-means clustering based on a prior selected number of clusters) and density-based clustering. The location importance in the research is based on four features: readings (time spent in a location), reading days (unique days when reading is present), visits (number of continuous sequences of visits), visit days (unique days when a visit is present).

Research in the travel domain was done on mining locations based on GPS trajectories for travel recommendations. The proposed locations are based on the number of total visitors, each user's previous travel experiences, and the correlation between locations. The locations retrieved from the GPS data can help users understand their surroundings and plan their travel routes [17].

A concept of a "personal" map was introduced that depicts a personalized and customized individual map [20]. A personal map can include but is not limited to individually

significant places and routes, shopping centers, workplaces, and paths. The research analyses GPS data and answer three questions; discriminating user's activities, predicting future movements and places, and inferring when users have broken their regular route. An approach for significant place extraction from raw GPS data used in the research is solely based on the individual's time spent at a particular location. Based on a set threshold, the significant places are filtered however, if the threshold of time spent is too high, then some significant places may be filtered out because they are visited often, but for short periods [20].

2.2 Mobile cloud application forensics

Mobile cloud computing utilizes the benefits of cloud computing, such as unlimited data storage and processing power, on mobile device applications. Google Maps application stores the majority of the data sources on cloud servers, with some data points stored on the device locally, including caches, application data, cookie preferences, and offline downloaded maps. All other data is stored in on-demand servers accessible to users through a Google account. This section reviews previous research done on mobile cloud application forensics practices and challenges.

The book “Digital Forensics: Threatscape and Best Practices” [6, Ch. 5] explicitly presents mobile device application forensic challenges and threats. Mobile phone applications are mainly installed by the user based on a specific need to utilize the app, excluding cases when the apps are included in the original installations, such as Google Maps for Android devices. It is of the highest importance for the mobile phone forensic investigator to gain knowledge of how and what data to extract from these different applications. All evidence discovered in the digital forensic process possesses unique characteristics and must be observed individually.

B. Martini, Q. Do and K.-K. Raymond Choo in [22] have analyzed seven Android cloud-based applications. The methodology used in the research was creating a copy of a physical image of the device and manually reviewing the files in the private app directory and on the mobile device's internal storage. The files that the authors believe hold a forensic interest linked with each application were outlined in the research.

N. Al-Mutawa, I. Baggili, and A. Marrington in [23] have presented a mobile forensic analysis of three social networking applications – Facebook, Twitter, and MySpace. The methodology used in the research consists of installing and executing various tasks on the smartphone applications and then acquiring a logical image of the device and performing manual forensic analysis. The outcome presents if the mobile device's internal memory stores certain activities from the applications. Depending on the device and operating system, data retrieved varies in provided value.

P. Sharma, D. Arora, and T. Sakthivel in [24] have developed and tested a forensic examination scenario of a cloud-based social networking application WeChat. Through five steps WeChat application was examined on both: the Android device and by tracing cloud artifacts. The forensic investigator extracted relevant log files from the traced cloud data and linked them with the data collected locally from the Android device. Both combined data sources improve the investigation accuracy. The overall suggestion is to include mobile cloud traceability in the mobile cloud forensic framework.

Similarly, H. Zhang, L. Chen, and Q. Liu in [25] have performed a complete mobile forensics analysis on four instant messaging applications on android devices. The devices are rooted, and the data stored locally on the device associated with the applications are manually analyzed and presented in the report.

Third-party application forensics on an iOS device was performed by A. Levinson, B. Stackpole, and D. Johnson in [26]. A forensic image of the mobile device was acquired, and data from third-party applications was manually analyzed to solve the designed situation. Various data points were discovered locally on the mobile device user data partition, including credentials, time-stamps, and geolocation data.

We observe that most research done in the mobile device cloud application forensics domain extracts the traces of data from the cloud-based applications stored on the mobile device's local memory. However, a vital aspect is recognizing that only a small part of the data is stored locally on the device, and the other parts are stored on the remote cloud storage accounts.

N. Samet, A. Ben Letaïfa, M. Hamdi, and S. Tabbane in [27] mention that mobile cloud computing solutions complicate the mobile forensic process because two different environments need to be considered separately: mobile device and cloud server. As

mentioned in their report, cloud environment forensics faces challenges connected with cloud virtualized architecture, lack of cloud-based forensics tools, and chain of custody preservation. They outline a significant issue of maintaining a chain of custody because data from the cloud can be accessed through any workstation irrespective of the device, and there are laws connected with accessing proprietary technology in a cloud environment.

S. Zargari and D. Benford have compared computer forensics and cloud forensics in [28]. In a cloud environment, it is impossible to secure, evaluate and document the crime scene. Data storage is on the cloud service provider's data center rather than in the evidence room, as in the computer forensic process. In cloud forensics, data is transported only electronically, compared to the physical transportation of a hard disk. Additionally, they state that recovering the deleted data in a cloud environment is more complicated. It is also more difficult to keep consistency in the data acquisition process. Acquisition time in cloud forensics is fast, however, computer forensics is generally a slower process. The same platforms are shared between different consumers as a result complicating the digital objects' ownership and the attribution of the data.

Almulla, Iraqi, and Jones [29] outlines the state of the art cloud forensics and arranges issues connected with cloud forensics in three categories: data and architecture-related technical issues and legal challenges. As mentioned, data preservation may become particularly difficult with current digital forensics practices and the dynamic environment of the cloud. The research also outlines that the focus of the reviewed solutions is on utilizing traditional digital forensic methods to analyze data stored in the cloud. However, there is a minimal focus on actually using cloud solutions to conduct digital forensics.

Several pieces of research were published to evaluate the challenges of mobile cloud forensics. We summarize the main challenges linked with cloud data as data attribution, data deletion, data validation, accuracy, and legal access to the cloud account. A critical observation [27] is presented that mobile cloud computing solutions must be considered as two separate instances of mobile device and a cloud server. To a great degree, mobile cloud application forensics research focuses on gathering the data linked with the applications stored locally on the mobile device. We observed a lack of publications on actually utilizing the cloud environment to retrieve the data rather than analyzing application data stored locally on the mobile device.

2.3 Mobile location data forensics

Location data and applications are discussed in [6, pp. 77–78]. Google Maps and Apple Maps can provide location-based information on traveled places and saved directions. It is mentioned that the data from these location-based applications can be retrieved from the app itself on a mobile device. Apart from location-based applications, other applications that utilize location data, such as Facebook and Twitter, can be used to approximate locations where posts or other activities are made.

M. Chernyshev, S. Zeadally, Z. Baig, and A. Woodward in [7] have performed a digital forensics examination of the suspect's location history from a mobile device. The evidence extraction scenario presents a case of performing level-3 imaging of the device following a standard procedure, including device rooting, image acquisition, and hash validation. Afterward, with the Autopsy tool, data is examined. The final location data was retrieved from camera photos in the image embedded geolocation metadata.

Moore, Baggili, and Breitinger [8] inform practitioners, researchers, and the forensic community on data extraction methods from mobile phone location-based applications: Google Maps, Apple Maps, Waze, MapQuest, Bing Maps, and ScoutGPS. After using .XRY software for data acquisition from the Google Maps application, the following files were discovered locally: destination history with start and end locations, search history, users preferences, account data, and other cloud data based on the user account. On the iOS device, only the user's last known location file was discovered. Comparing all analyzed applications Google Maps for Android was ranked as the third application based on the forensic value of the extracted data, proceeding with Waze and Scout. Google Maps for iOS was ranked 9th because it only stored the last location before the application was closed. Alternative research [26] on iOS device third party applications mentions that the Google Maps application stores bookmarks, driving directions, contact addresses, recent map searches, and information about the last located coordinate locally on the device. A.Edens in the Cell Phone Investigation series (2014) [30], states that digital forensic information retrieved from a Google account can exceed the amount of data retrieved from a mobile phone using forensic hardware or software programs.

Forensic analysis of smartphones reveals that geolocation data from a mobile device can be retrieved from Wi-Fi connections, location databases, and IP connections history. The

hotspot connection history collects and presents SSID (service set identified) and RSSI (received signal strength indicator) information, revealing where the mobile device last connected to the Wi-Fi. Additionally, IP connection history presents the mobile device's location through the collected data points such as MAC and VPN addresses, router IP addresses, and DNS names [31].

Forensic analysis of the Android TomTom navigation application was done by Le-Khac et al. in [9]. The methodology used was to physically extract, identify, decode, and analyze related files locally stored on the mobile device. Relevant files discovered after the extraction are application settings, destination history, saved favorite locations, and search history. The only method to see if a location was visited is if it is stored in the current GPS location file and if GPS coverage is available. It is also mentioned that the major challenge with location data extraction from mobile applications is the great diversity of the applications and the exclusive data storage formats from each source.

2.4 Forensically sound digital evidence

This chapter defines forensically sound evidence and the significance of using proper methodologies for the digital forensic process. Additionally, it justifies the importance of throughout Google Maps timeline examination and testing before it can be used as a data source for an expert report in court.

On October 27, 2017, a published article in “Times Union” newspaper states that Google location evidence in a criminal trial retrieved from the defendant's mobile phone was unsuitable for prosecution. The following statement was published to the prosecutors: “*failed to meet their burden of demonstrating that the science underlying Google location services has gained general acceptance in the relevant scientific community*” [2]. In a publicly known “Arizona case,” an innocent individual was arrested and retained in jail for one week in 2018, based solely on location history data provided by Google after a search warrant request. After receiving additional information, the suspect was released one week later, and no charges were pursued. However, the one-week arrest profoundly affected the arrested person's life and livelihood [32]. Case files in [33, p. 28] reference a murder trial where cell phone location was retrieved from the mobile device and cell tower data helping to reject a supposed alibi. The three publicly available case backgrounds present an issue and an opportunity that location data in mobile devices is

very valuable, however, if it is not tested in the scientific community, it can not be used on its own.

E. Casey and C. W. Rose in [33] discuss that forensic examiners should not rely on a single tool as it may significantly impede the digital evidence review and subsequent investigation. Gary Ernsdorff, a Washington State prosecutor, who has worked on several Google location data cases mentioned for the New York Times states: "*We are not going to charge anybody just because Google said they were there*" [32].

McKemmish in [34] has directly examined the definition and overall meaning of "forensically sound" evidence mentioning that the forensically sound process involves data preservation in its initial state and analysis without reducing the evidence value. He informs that there is an increasing trend to specify in the expert reports that the method and technology used are "forensically sound", indicating that the final result has not lost its evidentiary weight. Subsequently, evidence preservation in a forensically sound manner leads to admissibility in court. Eoghan Casey in [35] explains that admissibility is required by the court to identify if the presented evidence is accurate enough and will assist as a solid foundation for decision-making in the case. One of the primary issues that may prevent evidence from being admitted in court is improper handling.

S. Vömel and F. C. Freiling in [36] identify three criteria for evidence to be forensically sound: correctness, atomicity, and integrity. This definition was developed as a foundation for defining a forensically sound copy of physical memory. They also argue that forensically sound evidence may be altered to the degree that does not affect the reliability and authenticity of the data. The investigator can evaluate the soundness of the evidence and the acceptance criteria for each case separately.

Important factors supporting evidence authenticity are a chain of custody and integrity documentation. Chain of custody shows how the evidence was acquired, and integrity documentation supports the belief that the evidence has not been altered after acquiring [35]. Authentication of manual evidence is crucial to ensure that the data presents what the proponents claim to be. Chain of custody entails that the evidence is processed in its original form.

S. Goodison [37] presents a case where improper digital evidence handling has served as a factor to weaken a case of a murder trial. They also mention authentication and chain

of custody importance in the documentation of the digital evidence. Authentication establishes proof of who the digital evidence's factual owner is, and chain of custody is assurance that the preservation of the digital evidence is in its original form. [38] points toward the importance of reliability. For forensics findings to be admissible in court, forensic examiners must follow a well-defined process of data collection, analysis, and reporting.

R.B.Adams in his Doctoral Thesis work [39] presents a list of standards and guidelines required for forensic data acquisition. These are ISO standards, British Standards Institute (10008 standards), Association of Chief Police Officers Guide, International Organization on Computer Evidence, McKemmish rules, Gosh guidelines, and Brezinski and Killalea's guidelines. It is not easy to apply generally accepted standards to acquire evidence from rapidly changing and evolving applications.

S. Goodison in [37] display an issue related to the extremely high amount of data and skills required for digital evidence extraction. The backlog of data evidence for the examiners can be up to one year. Additional challenges mentioned in the workshop report are fast-changing technology and constraints in the budget for new equipment. The increasing complexity of acquiring data from Android mobile devices has also been mentioned in [40, p. 38] by A. Hoog. He states that even minor differences in Android versions must be extensively tested and validated prior they are perceived as forensically sound in the court.

To support the evidence, forensic examiners must hold a throughout understanding of the technology used. Presented evidence is often assessed to understand the real strength of the proposed hypothesis. In the "Handbook of Digital Forensics and Investigation" [33] it is mentioned that forensic examiners frequently use a scientific approach and conduct controlled experiments to gain insight into the program or the system. This approach is essential to minimize the tendency of leaning toward favoring a particular hypothesis.

DeMatteo, Fishel, and Tansey in [41] outline a rising concern that the court admits expert reports based on invalidated methodologies with a particular regularity. They explain that the evidence derived from invalid procedures has far-reaching consequences in court and that this type of expert evidence that does not have a solid scientific background can be called "junk science".

3 Research background

3.1 Toolkit

Four devices were used in the experiments (Table 2). Three were test devices, and one was used as a ground truth device to log the actual location.

Table 2. Technical documentation of mobile devices.

Model name	**Model number**	**Software**	**Type**
Galaxy S21 5G	SM-G991B/DS	Android 12.0	Validation
Huawei P20 Lite	ANE-LX2	Android 9.1.0	Test
iPhone 7 Plus	MNQM2ET/A	iOS 15.0.2	Test
Galaxy A3 2016	SM-A2310F	Android 7.0	Test

Ground truth data was captured using the application “GPS Logger Lite”. It is a free application that logs the device's location only based on GNSS signals. This application was selected because it is available to the author, and its accuracy is sufficient in the context of these experiments. The application, while running, detects the physical location of a device every second. The application is run on device Galaxy S21 5G (Table 3).

Table 3. Ground truth raw coordinate data extraction device.

Device name	**Software name**	**Location capture interval**
Galaxy S21 5G	GPS Logger Lite App (4.3.80)	1 sec

Based on a publication by Google, there are 13 different permissions types on an Android mobile device that can be adjusted based on the users' preferences. Google Maps application can use 8 of these permissions to provide the service. A complete list of available permissions with their practical usage is listed in Appendix 2. We allowed only location permissions for Google Maps (Table 4) on all test devices for this research. All other permissions were disabled and were not tested during the experimentation because we could not evaluate their impact on the outcome.

Table 4. Allowed permissions on the tested devices.

Name	Google Maps version	Allowed permission name
Huawei P20 Lite	11.15.3	Location
iPhone 7 Plus	6.3	Location (Always)
Galaxy A3 2016	6.1	Location (Always)

Improved location accuracy service (Figure 2) uses Wi-Fi, cellular networks, and other sensors concurrently to improve the final location. If Google's improved location accuracy service is turned off, it only uses the GNSS, which is proclaimed to be less accurate. Improve location accuracy service is enabled in all experiments.

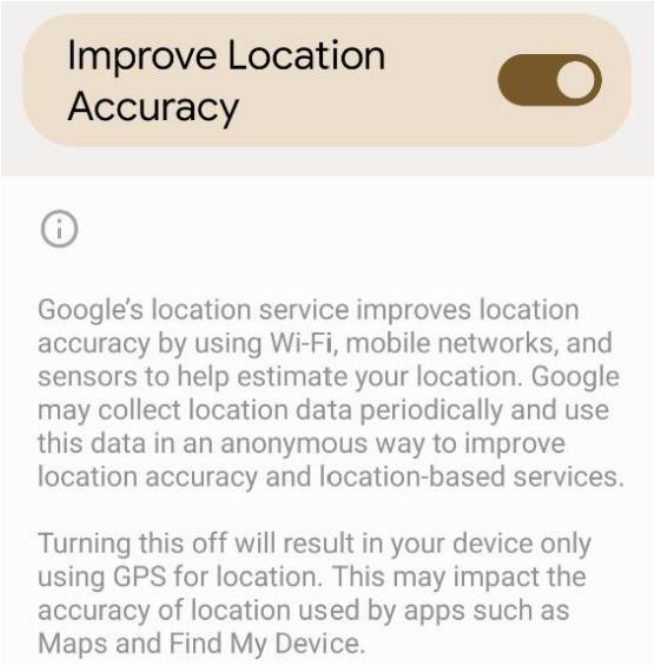


Figure 2. Improved location accuracy service.

Precise location service (Figure 3) within location app permissions allows the application to use the device's precise location. It is mentioned that if this setting is disabled, the application will only determine the device's approximate location. During the experiments, the precise location is turned on at all times.

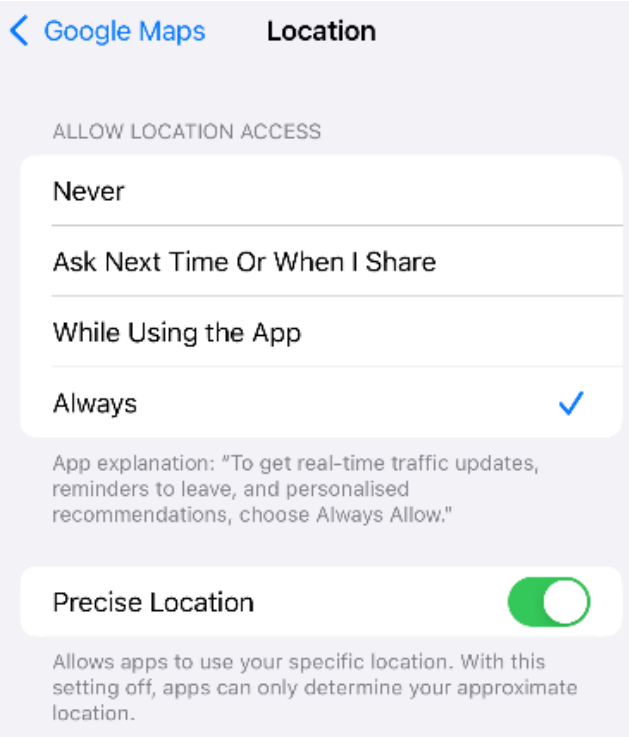


Figure 3. Precise location on iPhone device.

Advanced settings on mobile devices as presents in Figure 4 are Wi-Fi scanning and Bluetooth scanning. Wi-Fi scanning permission states that it improves the positioning accuracy for apps and services by scanning Wi-Fi networks, even if Wi-Fi is turned off on the device. Bluetooth scanning improves the positioning by scanning Bluetooth devices, even if Bluetooth is turned off.

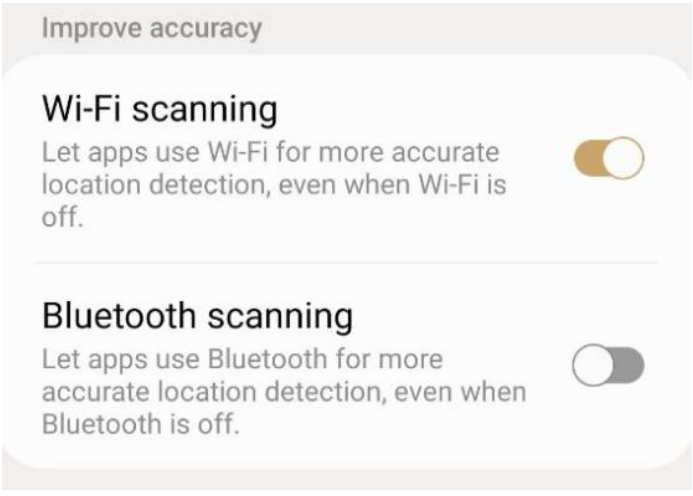


Figure 4. Wi-Fi and Bluetooth scanning on Android device.

Google account history settings personalization directly affects the performance and data extraction possibilities of Google maps. There are three history settings: Web & App activity, location history, and YouTube history. Web & App activity and location history

preferences must be turned on for the Google Maps timeline to generate any data. Each account used in the experiments has enabled Web & App Activity and location history data collection. An example of Google account history settings is exhibited in Figure 5.

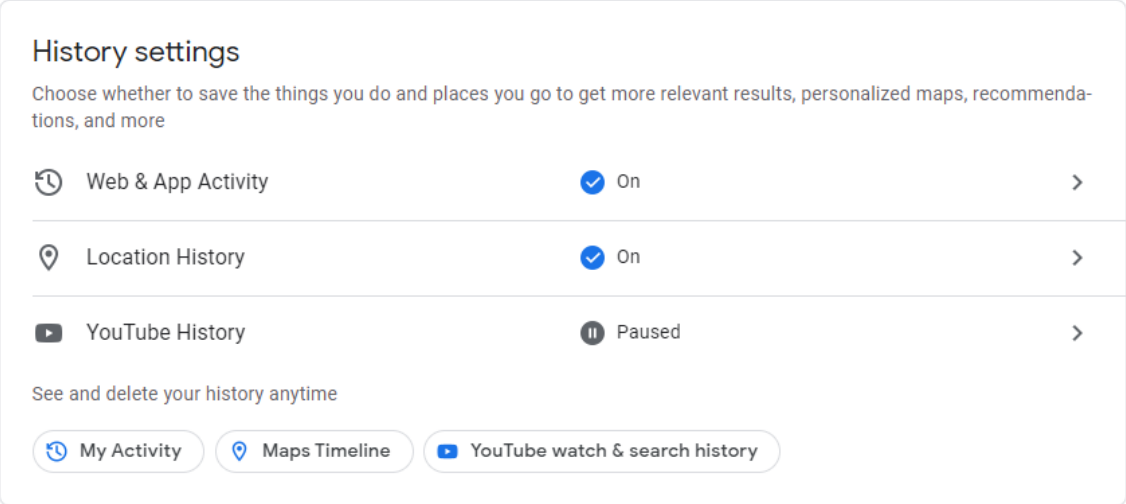


Figure 5. Google account settings.

Microsoft Power BI program is used as a data transformation and analysis tool. This program was selected due to its availability and ease of use. All transformations could be done in the transform data section, and tables could be joined clearly and understandably. Validation and test data tables are joined using one to many relationships (From test to validation) with a timestamp (DD/MM/YYYY HH:MM:SS) UTC+0 and device name index. The validation data table contains a physical location of every second of the experiment duration. If there are missing data, the previous location is repeated until a new physical location data point is received.

3.2 Analysis of location data from Google Takeout

Google Maps data was collected from 2 individual Google accounts with permanent residences in different countries. The data is summarized and presented in an aggregated form; therefore, no personal information is displayed. This section presents the general data structure and insight gathered from Google Takeout location data. From the analysis, we discover the main criteria tested in the experiments. In total, 557 000 raw location data points and 18 000 visited locations were analyzed.

3.2.1 Raw data records

Raw location data points are exported from Google Takeout as records.json files. All raw location history is exported in one file. Raw location data consists of seven values: latitude, longitude, accuracy, activity, source, device tag, and timestamp.

Each location data point is subject to an error rate presented as an accuracy feature in the dataset. The capture interval and accuracy of the received raw locations depend on the device's enabled navigation sensors and other external conditions. The primary location data sources in Google Maps are Wi-Fi, Cell, and GPS. A blank value indicates an iPhone device, as it does not save the source information. The Wi-Fi source is the primary navigation sensor providing the location data to the application. Two proceeding sources are cell towers and GPS. We assume that by disabling the Wi-Fi sensor on the device, the device's navigation accuracy will deteriorate. Number of collected raw location data points by location source are presented in Figure 6.

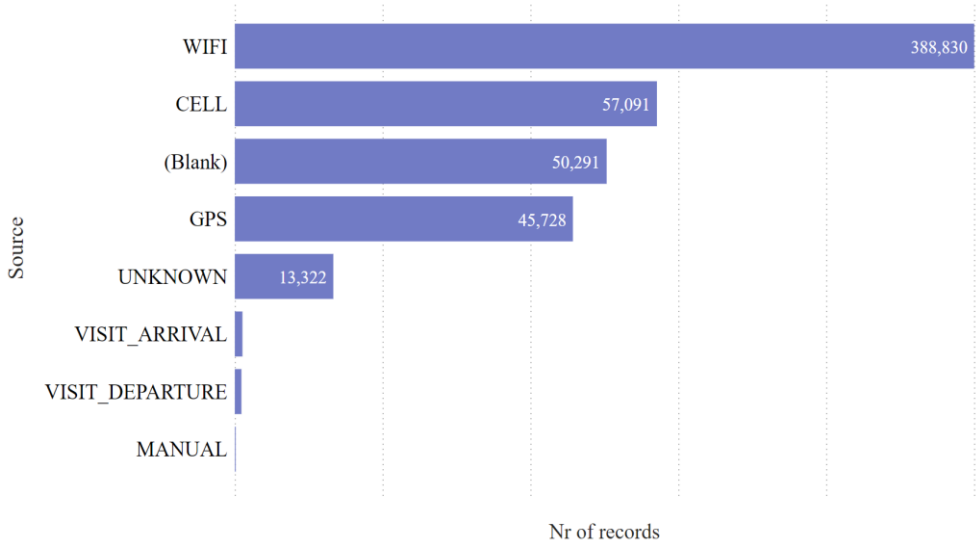


Figure 6. Number of records by source.

GPS, Wi-Fi, or CELL navigation sensors each hold a different error rate. We have aggregated all raw location point accuracies on the source level in Table 5. The table shows that the lowest median accuracy of 12 meters is for GPS source. 50% of all location accuracy values for GPS source are from 6 to 24 meters. For Wi-Fi source, 50% of values lay within the 20 to 27-meter range. Cell tower source holds the highest median accuracy of 800 meters, and 50% of values are within the 699 to 1899-meter range.

Table 5. Location accuracy by the source.

Source of location	Median of accuracy (m)	1st quartile of accuracy (m)	3rd quartile of accuracy (m)	IQR (m)
Cell	800	699	1 899	1 200
Visit_Departure	54	25	82	57
Visit_Arrival	48	35	66	31
Wi-Fi	23	20	27	7
Unknown	16	12	23	11
GPS	12	6	24	18

External circumstances also affect the location estimation techniques — for example, building structures, physical landmarks, and area density. Due to different navigation sensors and external factors, the accuracy varies from 3 to more than 1000 meters. As exhibited in Figure 7, the majority of raw location data points are focused around 3–100-meter accuracy values.

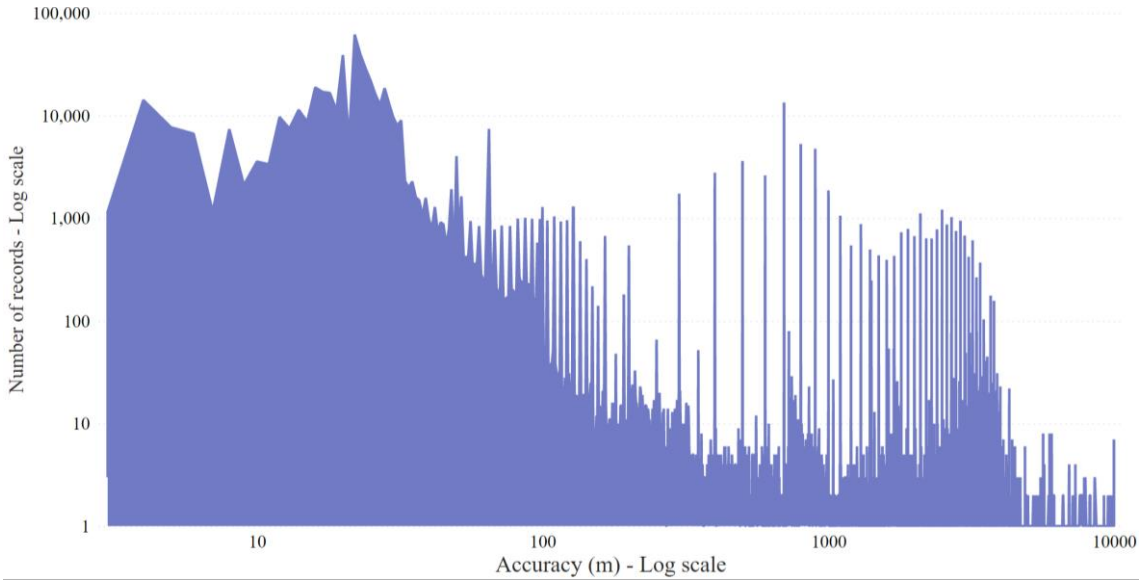


Figure 7. Number of records by the accuracy.

3.2.2 Semantic location history

Semantic location history consists of two separate groups of events: activity segments and place visits. Activity segments present the modes of travel and their relevant probabilities. Place visits present the possible visited locations and probabilities of these locations and places. A visit is an event that is displayed in the Google Maps timeline. If

the visit confidence is below a certain threshold, the visit will not be presented in the timeline. Each visit consists of relevant raw data points. Figure 8 presents a visit to "Pasadena Green Shopping Centre". The visit duration is 28.82 minutes, and within this timeframe, five raw data points were captured.

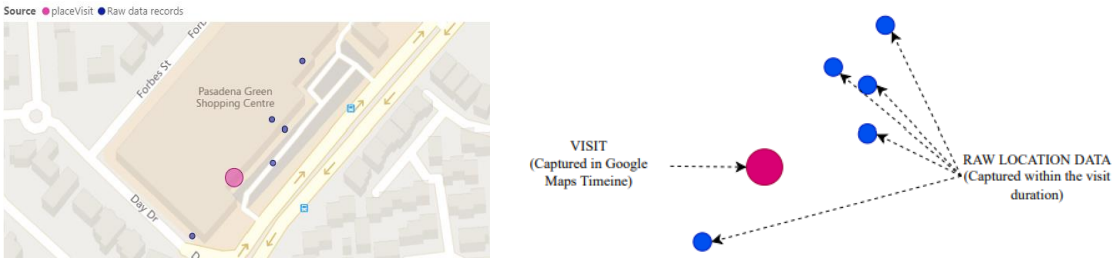


Figure 8. A visit with respective raw data points.

Place visit data consist of location, address, name, duration, Google placeId, latitude, longitude, other candidate locations, and edit confirmation status. The name presents the semantic name of the location. If the location's name is not available, then the address is presented as name.

Each visit contains a start time and end time. Start time is equal to the first raw data point timestamp within the visit, and end time is equal to the last data point timestamp for a visit. Visit duration is based on the timestamps of the raw data locations. The visit starts when the device has been present in one place for a specific period. The visit ends when the distance between the visit and the current raw data location exceeds a certain threshold, indicating that the device has moved away from the location.

Figure 9 depicts the haversine distance between the current raw data point and the previous raw data point. Number 1 indicates the first raw data point of the visit. If the visit has four raw data points, the average distance within the visit is calculated between numbers 1;2;3;4. Consequently, the first raw data point depicted as 1 is compared to the previous raw data point collected by the device. The second raw data point, depicted as 2, is compared to point 1. As projected, we observe that the average distances within the visit are significantly lower than the distance between the first data point of the visit and the previous raw data collected by the device.

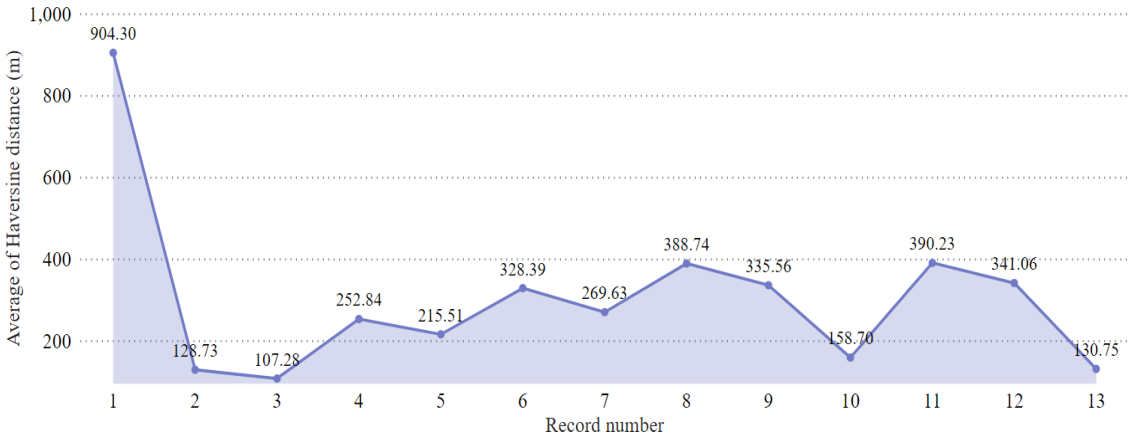


Figure 9. Haversine distance between the current and previous raw data point.

Each visit holds location confidence and visit confidence that varies depending on several factors, including the confidence of the raw location readings of the device. There are three place confidence intervals- high, medium and low confidence. These intervals inform the user about the certainty of the visited place. The visit confidence value indicates if the visit will appear in the Google Maps timeline. For a visit shown in the Google Maps timeline, the confidence is never below 60 points. The visit confidence is generally unchanged for each high, medium, and low interval. Figure 10 displays that the highest location confidence is for high confidence visits, following medium and low. For high confidence visits 50% of the location confidence values ranges between 80 to 83.

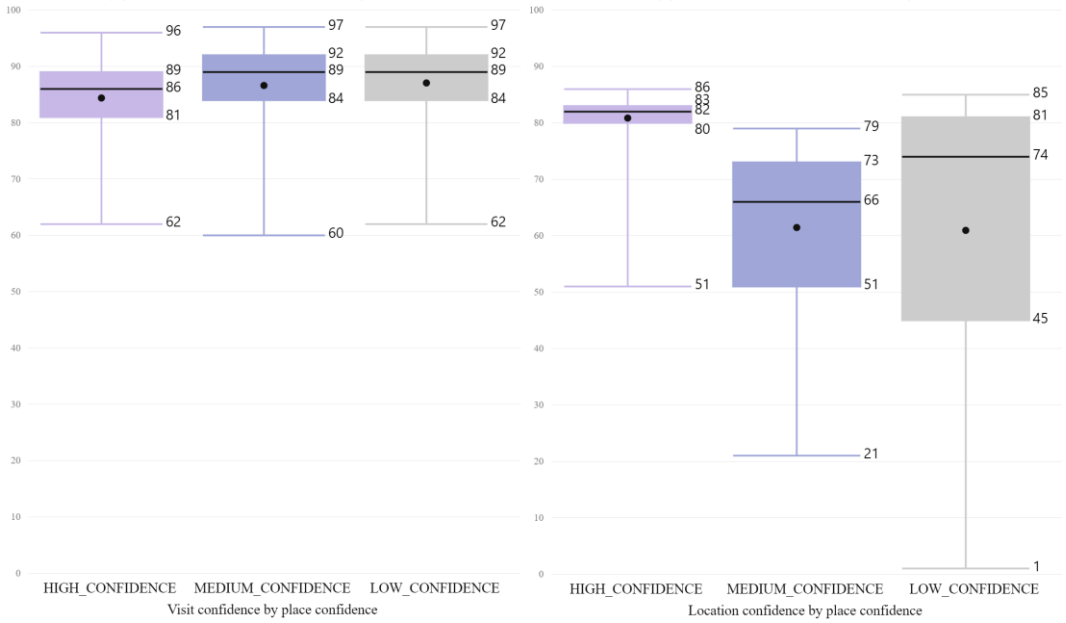


Figure 10. Average visit and location confidence by the place confidence interval.

There is a minimum time duration for a visit to be presented in the Google Maps timeline. For the reviewed dataset, the minimum duration of a visit is 2 minutes. The measure

indicates the minimum duration a device must be in a single place for a visit presented in the Google Maps timeline. The minimum visit duration for high confidence visits is 5 minutes that is above the minimum durations for medium and low confidence visits. Minimum duration of the visit by place confidence is presented in Figure 11.

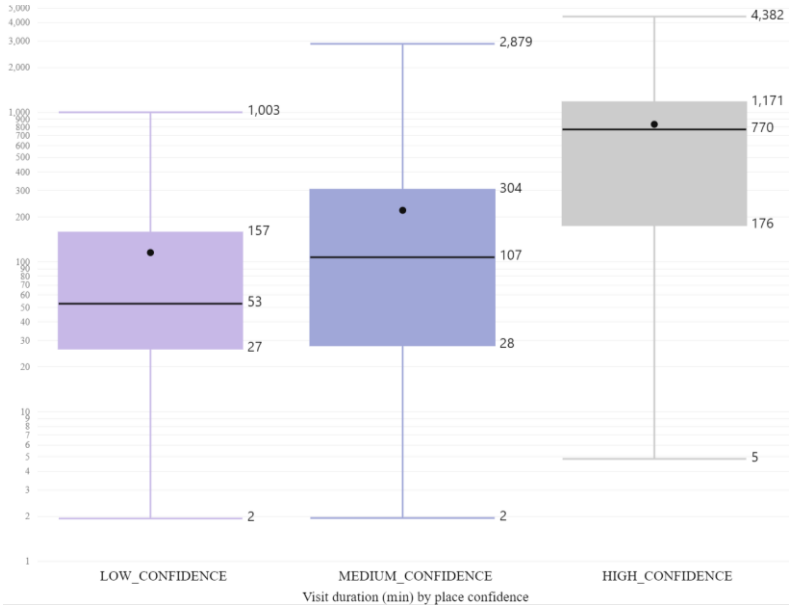
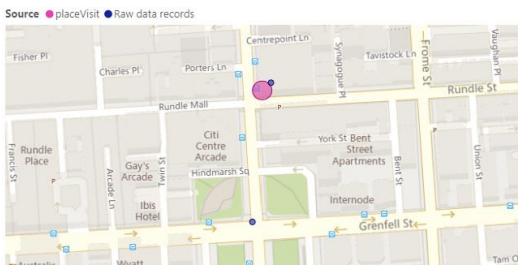
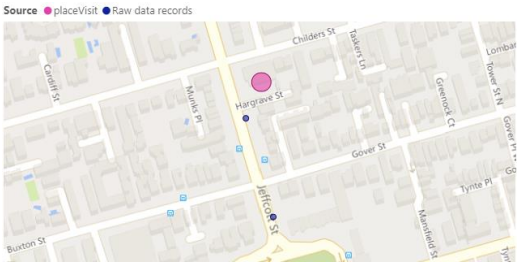
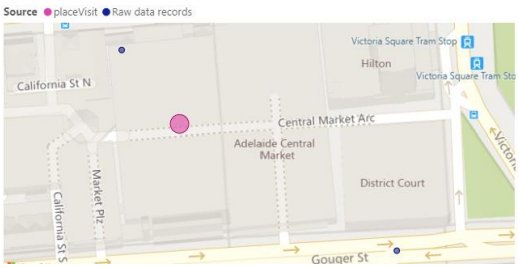


Figure 11. Visit duration (min) by place confidence.

Table 6 presents three examples where a visit is detected with only two raw data points and a duration of 1.93, 1.95, and 2.25 minutes.

Table 6. Minimum duration visit examples.

Visit of building Centrepoint with the duration of **1.93** minutes and a confidence of 75. The haversine distance between raw data point and place visit location is 162m. Records accuracy proposed by Google is 14m.	
Visit of Estonian Cultural center with the duration of **1.95** minutes and a confidence of 65. The haversine distance between raw data point and place visit location is 48m. Records accuracy reported by Google is 24m.	

Visit of Adelaide central market with the duration of **2.25** minutes and a confidence of 71. The haversine distance between raw data points and place visit location is 155m. Records accuracy reported by Google is 4m.	

Two variables were used to characterize outlier visits: Google reported accuracy of the raw data point and the haversine distance between the raw data point and the visit location, as presented in Figure 12.

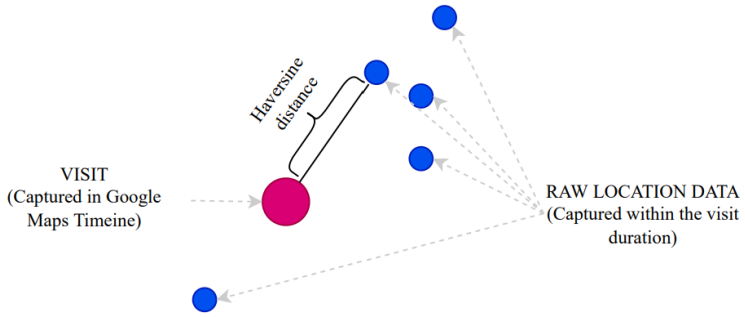


Figure 12. Visit and raw data points.

High confidence visits have the mean of google recorded accuracy of 34 meters and mean of haversine distance of 46 meters. Medium confidence visits have the mean of google accuracy of 51 meters and haversine distance of 142 meters, low confidence of 88 meters and 142 meters, respectively. The analysis suggests that the visit confidence is significantly dependent on Google's accuracy estimate and the distance between the visit location and respective raw location data points. Figure 13 displays the haversine distance and Google's accuracy by place confidence.

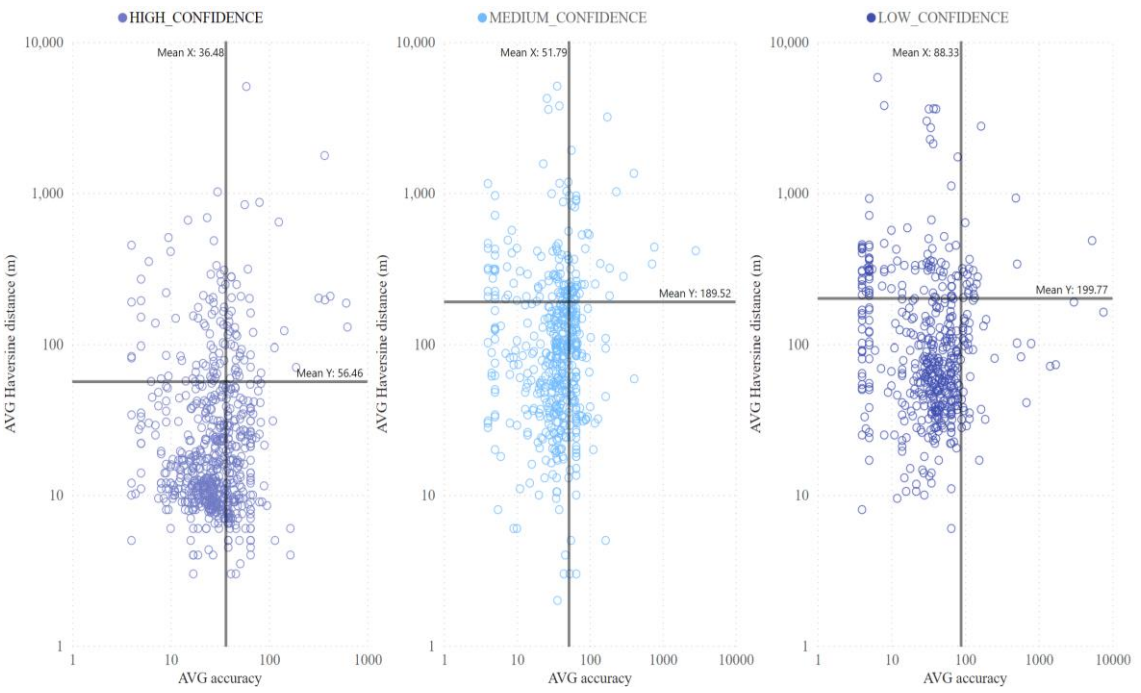


Figure 13. Accuracy and haversine distance of visits.

There are four summarized findings learned from the raw data analysis. First, the accuracy of raw data points captured by Google maps is essential to identify place visits correctly. If the accuracy of raw data points is unreliable, the timeline and visited places will also be unreliable. Second, the duration of the visit must exceed a threshold for the visit to be presented in the timeline. Depending on the place type, the threshold may vary. Third, the distance between the place-centered location and raw data points is lowest for high confidence places and increases for low confidence places. Fourth, the average distance within the visit is significantly lower than the distance between the first raw data point of the visit and the previous raw data point captured by the device.

3.3 Navigation sensor accuracy experiment

The Google Maps timeline for each account is based on several variables, and an important variable is the raw location data point accuracy received and stored by the mobile device and Google account. Before analyzing Google Maps timeline visit identification accuracy, we will look at the accuracy of the raw location data points received by the devices and accounts. There are three questions we will answer through the coordinate drift experiments.

1. How often is data captured in the test dataset under different variables?

We assume that the raw location data capture interval is dependent on the device navigation sensors. With this value, we present with what navigation sensors the location is captured most frequently. With an increased number of raw data points, there is a higher probability that the semantic location will be presented correctly in the Google Maps timeline. A lower interval of raw data points allows the Google Maps timeline to adjust more precise route mapping and visit duration calculation.

Equation 1 Raw data capture interval

$$\frac{experiment\ time\ (min)}{nr.of\ captured\ data\ points} = capture\ interval$$

2. What is the average accuracy proposed by Google for the raw data points?

Google services present the accuracy of each raw data point captured. Average accuracy presents the assumption made by the Google algorithm on the level of accuracy captured by the device and the service. With a different variation of device settings, we will validate the dependencies between the settings and accuracy presented by Google.

Equation 2 Google Maps average accuracy

$$\frac{Google\ accuracy}{number\ of\ records} = average\ accuracy$$

3. What is the haversine distance between the raw data points and validation dataset?

The approximate shape of the Earth is an oblate spheroid. The haversine formula calculates the distance between latitude and longitude points on a spheroid shape. We will present the shortest path between the two raw data points by applying the haversine formula to the test and validation datasets' latitude and longitude points. The shortest path represents the shortest distance between the test and validation data points. We assume that increased accuracy of the raw location data points increases the accuracy of semantic locations presented in the Google Maps timeline.

Equation 3 Haversine distance between test and validation data

$$hav(\theta) = hav(\varphi_2 - \varphi_1) + \cos(\varphi_1) \cdot \cos(\varphi_2) \cdot hav(\lambda_2 - \lambda_1)$$

$$\text{hav}(\theta) = sin^2(\theta/2)$$

$$D = 2r \cdot \sin^{-1}\left(\sqrt{\sin(\frac{\varphi_2 - \varphi_1}{2})^2 + \sin(\frac{\lambda_2 - \lambda_1}{2})^2 \cdot \cos(\lambda_1) \cdot \cos(\lambda_2)}\right)$$

φ_1 – latitude of validation

φ_2 – latitude of test

λ_1 – longitude of validation

λ_2 – longitude of test

r – radius of a sphere

D – distance between two locations

We present haversine distance in meters between test and validation data sets. The radius of the Earth is 6 371 * 1 000 meters, used as the radius of a sphere for haversine distance.

3.3.1 Experiment organization

The experiments are conducted on three devices and solely on foot. We select a specific route in the high-density city of Riga, Latvia. The following route has been developed in the Google Maps application, and it will be followed during the experimentation. The Google Maps application states that the entire route is 4.5 kilometers and will take approximately 56 minutes of walking time. The route is exhibited in Figure 14.

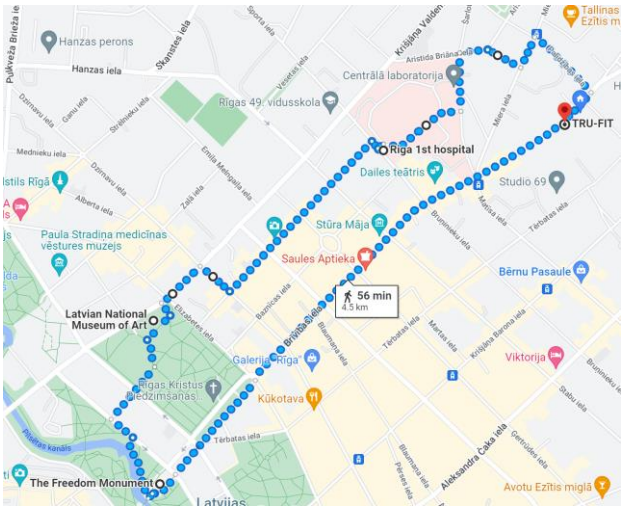


Figure 14. Route of navigation sensor impact experiment in Riga, Latvia.

The research [3] performed by the Netherlands forensic institute tested Google Maps timeline accuracy based on the phone configurations: GPS, Wi-Fi, 2G, and 3G networks. We approach the research question similarly; however, the tested navigation sensors on the devices are Wi-Fi + Bluetooth, location services and mobile networks (Table 7), because mobile device users primarily enable or disable these navigation sensors.

Table 7. Tested device navigation configuration.

Phone configuration type	**State**
Wi-Fi and Wi-Fi scanning (let apps use Wi-Fi for more accurate location detection, even when Wi-Fi is off)	ON/OFF
Location services (GNSS)	ON
Mobile networks (cellular)	ON/OFF
Bluetooth and Bluetooth scanning (let apps use Bluetooth for more accurate detection, even when Bluetooth is off)	ON/OFF

Each experiment iteration follows a specific procedure. Firstly, all devices are loaded to 100% battery level. This detail is important because a low battery level and a power saver mode can affect the performance of satellite or network-based navigation, and we want to exclude battery level effects on the outcome. Afterward, we turn all test devices on and sign-in to all sock puppet Google accounts. Depending on the experiment, relevant settings are turned on or off.

Google Maps application on a mobile device requests and collects location information only when location services are turned on. Therefore, location services are turned on in all experiments. Location services allow using any satellite constellation within GNSS to position the device. We execute six experiments with various device configurations to assess the timeline performance in different conditions. We start the experiments by analyzing the performance when all device configurations are enabled. Afterward, we disabled Wi-Fi, mobile network, and Bluetooth, with the last configuration only enabling location services. The experiment device configuration matrix in Table 8 presents the device conditions that are tested.

Table 8. Coordinate drift experiment setup matrix.

Experiment Nr.	Location services (GNSS)	Wi-Fi	Wi-Fi Scanning	Mobile network	Bluetooth	Bluetooth Scanning
1.	ON	ON	ON	ON	ON	ON
2.	ON	ON	ON		ON	ON
3.	ON			ON		
4.	ON					

After phone settings are configured, we go to the start position of the route and turn on the GPS Logger Lite Application. It is critical to verify that the application that collects ground truth data works and collects physical location every second. An example of a functioning GPS Logger lite application is presented in Figure 15.

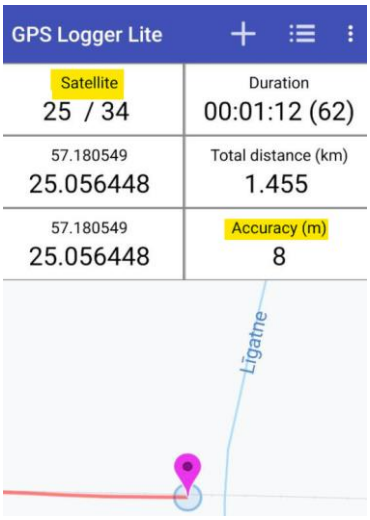


Figure 15. GPS Logger Lite.

After the technical setup is established and validation data is being collected, we follow the specified route with all the devices. After reaching the end destination, we turn off the "GPS Logger Lite" application and collect the data. Data collection is performed after each iteration from Google Takeout and stored on a Microsoft Surface Pro 6 computer. Validation data is downloaded from GPS Logger Lite application on the Samsung S21 device and uploaded to the computer.

3.4 Visit detection accuracy experiment

Based on the analysis of Google takeout data, we formed experiments to test the probability for different types of visits to be presented in the Google Maps timeline. The experiments will present if certain place types hold a higher probability to be presented in the Google Maps timeline as visits.

Similarly, as in the research "Accuracy and privacy aspects in free online reverse geocoding services" [13], we will apply weights for each visited location to calculate the average likelihood that the location will be presented in the Google Maps timeline. If the most exact location is presented in the timeline, we apply 1; if a location is near the actual visited location, we apply 0.5; if the location is not present, we apply 0. The final result is presented as a success rate in percentage for a location, device, and type of the visit.

3.4.1 Experiment organization

Experiments are conducted in two iterations. Dwell time for each visit is 10 minutes. Ten different locations are selected in a small town - Balvi, Latvia, and the other ten locations are in Riga, Latvia. Two different size cities were selected to include also rural locations. There are three location data types: semantic, logical, and physical locations. All three location data types are presented in this example: a physical location latitude longitude points 56.960434, 24.124547 after reverse geocoding is Bruņinieku Street 5, Riga, LV-1001 that holds a semantic meaning of Riga 1st hospital. All location types are reviewed in the experiments. Visited locations in Balvi, Latvia are presented in Table 9.

Table 9. Visited locations in Balvi, Latvia.

Nr.	**Type**	**Physical location (Lat/Long)**	**Full address**	**Semantic meaning**
1	Rural	57.125228 27.212260	-	-
2	Rural	57.125161 27.224335	-	-
3	Rural	57.124963 27.230343	Tala street 2, Balvi parish LV-4501	-
4	Town	57.132366 27.269472	Brivibas street 65, Balvi LV-4501	Labais shop
5	Town	57.133360 27.267442	Brivibas street 61, Balvi LV-4501	Balvi bus stop
6	Town	57.132474 27.263319	Partizanu street 8, Balvi LV-4501	Zebra pizza restaurant
7	Town	57.138996 27.257718	Krasta street 3, Balvi, LV-4501	Virsi gas station

Nr.	Type	Physical location (Lat/Long)	Full address	Semantic meaning
8	Rural	57.143838 27.232001	Stacijas street 25, Kubulu parish LV-4501	-
9	Rural	57.144711 27.249415	-	-
10	Rural	57.129364 27.243814	7 Ezermalas line 79, Balvi LV-4501	-

Three visits (1;2;9) do not have an address connected with the physical location, three visits (3;8;10) have only an address without semantic meaning, and four visits (4;5;6;7) have an address and a semantic meaning linked to it. All locations visited in Riga, Latvia are presented in Table 10 and they have a full address or a full address with a semantic meaning.

Table 10. Visited locations in Riga, Latvia.

Nr.	Type of visit	Physical location (Lat/Long)	Full address	Name
1	Outside	56.960292 24.130988	Brivibas street 90, Riga LV-1001	Narvesen shop
2	Outside	56.959685 24.126893	Brivibas street 75, Riga LV-1001	Dailes theatre
3	Outside	56.958246 24.125402	Brivibas street 76, Riga LV-1010	Caffeine cafe
4	Inside	56.956432 24.122067	Brivibas street 58, Riga LV-1011	Drogas shop
5	Inside	56.951966 24.114872	Brivibas street 30, Riga LV-1050	Street Food Point cafe
6	Inside	56.950500 24.111628	Z. A. Meierovica Boulevard 18, Riga LV-1050	McDonald's
7	Outside	56.954109 24.111510	Kalpaka Boulevard 6, Riga LV-1050	-
8	Outside	56.957499 24.117591	Skolas street 11, Riga LV-1010	-
9	Outside	56.960434 24.124547	Bruninieku street 5, Riga LV-1001	Riga 1st hospital
10	Inside	56.958548 24.125503	Brīvības street 78, Riga LV-1001	Maxima X grocery store

All experiment iterations are performed on three test and one validation device. All navigation sensors are enabled on the test devices. Validation data is collected throughout the experiment, including the time between changing the locations. The experiment starts in the first location. GPS Logger application is turned on. After visiting all locations, the

GPS Logger application is turned off. Afterward, data is extracted, transformed, and loaded on Microsoft Surface Pro 6 computer. Microsoft Power BI software is used for data analysis and presentation.

3.5 Mobile device navigation techniques

User location timeline accuracy depends on the specific device's navigation techniques. Google Maps timeline output is primarily based on the physical location points and secondly on the algorithms applied to these points. As our research is focused on outdoor localization, we review three main outdoor localization techniques used on mobile devices: GNSS (Global Navigation Satellite System), WLAN, and Cellular network. Apple mobile devices can use only one of three technologies to access geographic locations: GNSS, cellular network, and Skyhook Wireless Technologies based on wireless access point locations [26]. These techniques are used in the research as modifiable variables; therefore, understanding each technique is crucial. Indoor localization techniques are extensively reviewed in [42], [43], and [44]. A. Alarifi *et al.* in [42] outline indoor location estimation techniques that do not require specific hardware and utilize buildings infrastructures such as WLAN, Mobile cellular network, and Bluetooth.

3.5.1 Global Navigation Satellite System

A global navigation satellite system (GNSS) is satellite constellation that provides three services: positioning, timing, and navigation [45]. The satellite navigation system provides unlimited number of users with the receiving equipment three-dimensional position, velocity and time disseminated within the Universal Coordinated Time (UTC) [46]. The system consists of satellites that are orbiting the Earth and using trilateration method calculating the location of an object [47]. GNSS utilizes Time of Arrival (ToA) measurement on several satellites simultaneously to determine the three-dimensional position of an object. The distance from the emitter to the receiver is calculated by multiplying the speed of signal versus the propagation time [46]. Flaws of any navigation positioning system arises in indoor settings and highly dense areas when the signal is blocked [48].

There are various GNSS services designed by different regions and countries. The major navigation systems are Global Positioning System (GPS) – North America, BeiDou

Navigation Satellite System (BDS) – Peoples Republic of China, Galileo – European Union, GLONASS – Russian Federation, Quasi-Zenith Satellite System (QZSS) - Japan and Indian Regional Navigation Satellite System (IRNSS) – India [45]. A global positioning system (GPS) is one of the most widespread systems. It is defined as "a system by which signals are sent from satellites to a special device, used to show the position of a person or thing on the surface of the earth very accurately" [49]. The widespread usage of GPS outdoor localization techniques is due to practical accuracy, availability, and reliability. U.S Government states that SPS sent signal accuracy 95% of the time is below 2m with actual performance below 0.643m on 20th April 2021 [50].

3.5.2 WLAN based localization

IEEE 802.11 standard is more generally denoted as WLAN or Wi-Fi, and it is the most widely used standard for wireless communication between devices [43]. A normal WLAN coverage range is 50m – 100m [44]. The location of a device connected to a private or public wireless access point can be estimated. Several techniques to estimate the device location using its connection to WLAN are RSSI (Received Signal Strength Indicator), TDOA (Time Difference of Arrival), TOA (Time of Arrival), (AOA) Angle of Arrival, RTOF (Received Time of Flight) [44]. Received Signal Strength is the most common WLAN positioning technique because it can be integrated without any additional hardware modification on any device integrated with Wi-Fi [51]. RSS also does not hold the complexity of angular measurements and time delay issues opposite to the other techniques [43]. If the received signal strength becomes weak, the device can gather location information from other navigation sensors.

3.5.3 Mobile cellular network localization

The mobile cellular network is also denoted as GSM (Global System of Mobile Communication), consisting of many base stations and respective signal areas [52]. Base stations cover an area by their radio signal, creating a cellular network coverage. The mobile signal strength is affected by distance from the base tower, physical obstructions, building structures, physical landmarks, network congestion, and interference [52]. In rural areas, the radio transmitters are configured to cover large areas, whereas in urban, smaller radio transmitters are used to cover limited areas but provide high capacity. Based on the selected geographical location, the typical coverage radius varies from 200m for Picocell base stations for high rise buildings, 1-2km for Microcell base stations in urban

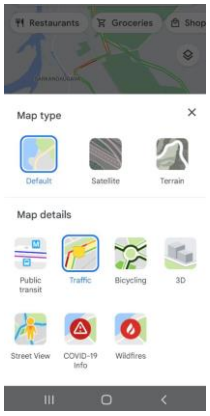
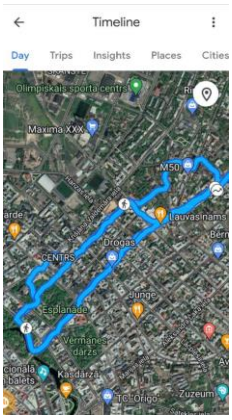
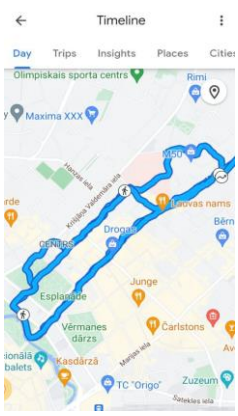
locations, 5-32km for Macrocell in suburban areas, and 50-150km for Macrocell Extended Reach in rural areas [52]. Depending on the cell size, the accuracy of the location varies between 50-200m [44]. Paper "Mobile Location Estimation in GSM/UMTS" [53] in detail explains all location estimation techniques within the GSM network, such as RSSI (Received Signal Strength Indicator), AOA, TOA. Paper "Outdoor Location of Mobile Devices Using Trilateration Algorithms for Emergency Services" [54] explains Cell of Origin, AoA, and Trilateration location estimation techniques. GSM location estimation of a mobile device is mainly used for outdoor environments. Indoor location can be estimated if the building has strong RSSI [44].

3.6 Geographic Information Systems

A Geographic Information System (GIS) presents geographic data on the Earth's surface. These systems allow us to understand and use geographic data in various applications. With the fast-paced development of mobile devices, an active research field has appeared called Mobile GIS. It unites geographic data transformation, management, and visualization to the end user. Mobile GIS based on location-based services is a system combining GIS, GNSS, internet, and mobile communication technologies [55]. In this research context, the GIS system refers to the Google Maps application.

In 2005 a design idea was presented in the IEEE International Geoscience and Remote Sensing Symposium to combine Mobile GIS and location-based services. The proposed system architecture comprises the client and the server connected wirelessly through a network [55]. The design idea is similar to today's Mobile GIS applications, including Google Maps. Different geographic representations of locations on a surface of the Earth are denoted as thematic layers. D. Arctur and M. Zeiler mention that thematic layers are assembled of geographic elements in order for the location points to hold a logical representation and insight of the raw data [56, p. 4]. Various thematic layers are overlayed to present one integrated map with information necessary to the user. For example, a road transportation network is represented by several layers, such as streets, intersections, urban areas, and bridges [56, p. 4]. Google Maps allows users to select from three map types: default, satellite, and terrain (Table 11). Google Maps timeline can also be presented in these three map types. Additional details can be selected to overlay the map, such as public transport, traffic, or bicycling.

Table 11. Google Maps available map types and map details.

Google Maps type selection	Google Maps timeline Map type: satellite	Google Maps timeline Map type: terrain
		

The development of geographic information system integration in mobile technologies was foundation for further advancements in location-based services [57].

3.6.1 Location-based services

The beginning of location-based services dates back to 1996, when the United States government passed a mandate for mobile-network operators to locate emergency callers with a certain accuracy level [58]. Since then, the emergence of low-power positioning techniques (satellite navigation and network-based) have endorsed a rapid change in location-based services' nature and widespread usage. The ability to wirelessly transmit spatial data from a mobile device is the foundational requirement for location-based services to work [55].

Location-based services are constructed from four key components: mobile device, positioning, communication network, and service and content provider [59]. Positioning is the selected localization technique for the device to determine the physical location. A communication network transfers the information between the mobile device and the service provider, and the content provider models the received information based on specified algorithms. The three principal components combined are responsible for the main tasks of location-based services: positioning, data modeling, and information communication [59]. Data modeling is the component where the context factors are taken into consideration. Referring to research on Mobile GIS based on LBS [55], the technical

design of LBS comprises a mobile device that sends its precise location to a service center, queries a database, and retrieves information based on the initial location.

Research done by H. Huang, G. Gartner, J. M. Krisp, M. Raubal, and N. Van de Weghe have combined location-based services' main trends and research agenda in [60]. They mention that location-based services should not solely rely on the location because each user's output is specific. The context factors affect the output that is provided. Depending on the specific service, different context factors may be relevant. Several factors that may be considered are physical surroundings, system properties, time, mobile map user, location, navigation history, and orientation.

P. Bellavista, A. Küpper, and S. Helal in [58] state that LBS have switched from reactive to proactive by automatically adjusting the output based on predefined changes in the user's movement. Additionally, LBS are application-oriented rather than only providing the location information after a request and are part of dynamic content presented to the user through user interface based on the current user's location and the execution context.

Explicit definitions and differences of location-aware systems and location-based services are outlined by H. R. Schmidtke in [61]. A location-aware system is defined as one that provides services based on the user's current location. However, location-based services deliver output that contains information about the user's location history. Contrary to location-based services, the main difference between the two notions is that location-aware systems do not require users to share and collect their location information.

3.6.2 Google Maps

Google Maps is an application and web platform owned by Google Inc. The application is pre-installed on most Android devices and widely used in other operating systems. The service is free of charge and can be used with and without an active Google user account. It provides navigation, traffic, street view, location search, and satellite imagery. The architecture of Google Maps is outlined by P. Gilski and J. Stefański in the “Survey of Radio Navigation Systems” [62]. The architecture is presented in Figure 16.

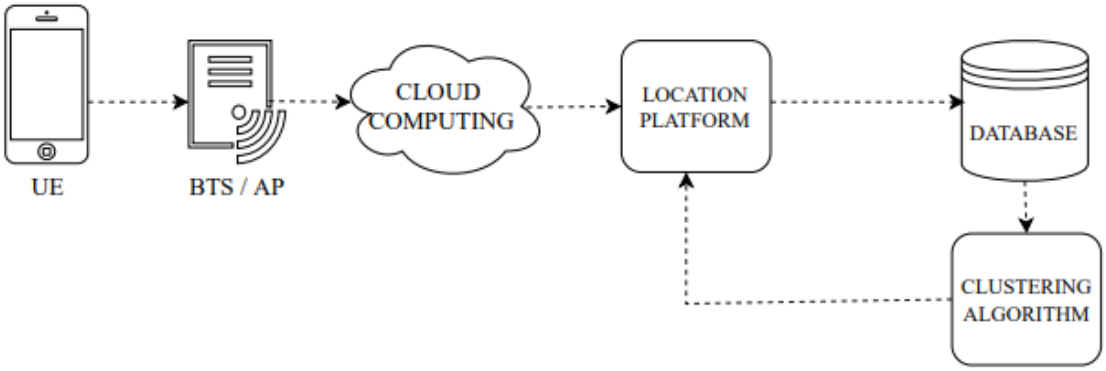


Figure 16. Architecture of Google Maps [62].

The user equipment (for example, a mobile device) is linked with a base transceiver station or another access point connecting it to a network. Through the network, the location platform (Google Maps) sends the location data to the database, where it is stored and processed. When requested, the location platform retrieves the transformed information and presents it to the user.

Google Maps timeline is a service within the Google Maps application. The application's principal purpose is to present the places and routes the user may have taken based on collected and processed raw data points. The timeline can be edited and deleted by the user. The service is only available for users who have signed in to their Google accounts and have enabled location services. The ability to collect and store history data is integral for the application to perform; therefore, users must also enable web & app activity data collection.

Google Maps timeline provides its interpretation of the place symbolic names that the user visits. Based on the user's profile and physical location data history, a place name is selected from a list of possible locations in the dictionary. The recommended place can be approved or declined and changed by the user. Google states that any searches initiated by the user can affect the visit recommendations in the timeline [63]. Although there is no readily available data to estimate how frequently this occurs in practice. The Google Maps timeline provides additional variables to the location history. As mentioned by Google policies, the location is not only based on the device's geographical information. To improve the contextualization of the data, Google also uses past activity and labeled place analysis concurrently with the device location data [64].

4 Results

4.1 Navigation sensor impact on data

4.1.1 Experiment results

Experiment 1. presents the performance of the Google Maps timeline when all device settings that directly impact location data are enabled. Device configuration in experiment 1 is presented in Table 12.

Table 12. Experiment 1 device configuration.

Experiment Nr.	Location services	Wi-Fi	Wi-Fi Scanning	Mobile network	Bluetooth	Bluetooth Scanning
1.	ON	ON	ON	ON	ON	ON

The experiment duration was 64 minutes. The performance of Android devices is better than iOS, with an average data capture interval of 2.21 for Huawei P20 Lite and 1.94 for Samsung A3. iPhone 7 collected 11 data points leading to the capture interval of 5.82. The average Google provided accuracy for raw data points is 29.40 meters. For the same raw data points, average Haversine distance to the ground truth location is 22.09 meters. Figure 17 exhibits experiment 1 results. The two measurements indicate that Google is less optimistic when calculating the accuracy and the actual accuracy is more precise.

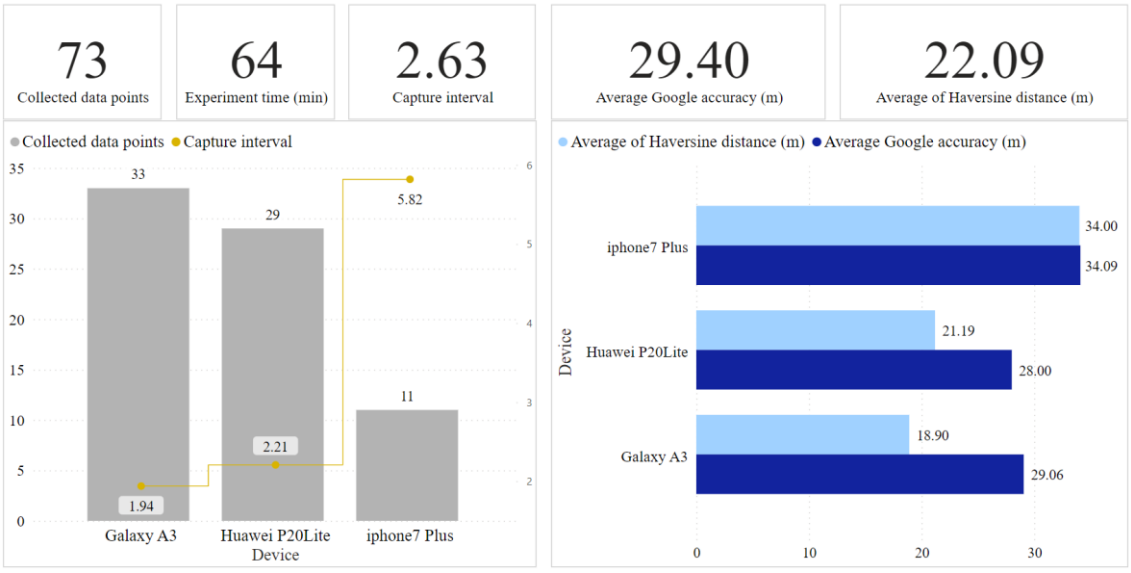
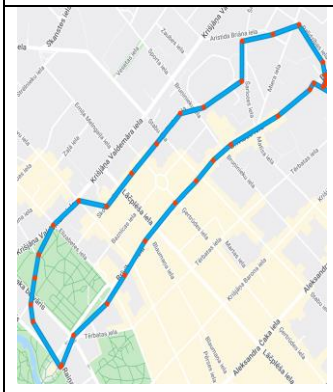
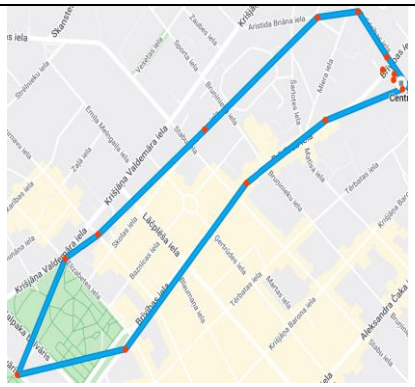
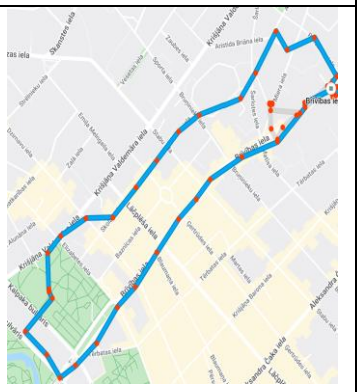


Figure 17. Experiment 1: Raw data point analysis.

The number of collected data points, capture interval and accuracy directly affect the timeline presented in the mobile application. The timelines in Table 13 are created based on the red data points. A decreasing number of red data points lead to a less precise timeline. We observe that the visual presentation of the route is worse for iPhone 7 Plus, because of significantly higher sample interval. Both timelines presented on the android devices depict all streets, crossings, and major turns.

Table 13. Experiment 1: Google Maps timeline.

Huawei P20 Lite	**iPhone 7**	**Samsung A3**
		

The collected validation and test data points are presented in Figure 18. We observe a pattern of raw data point collection – the raw data points are collected at the same places and with a certain consistency.

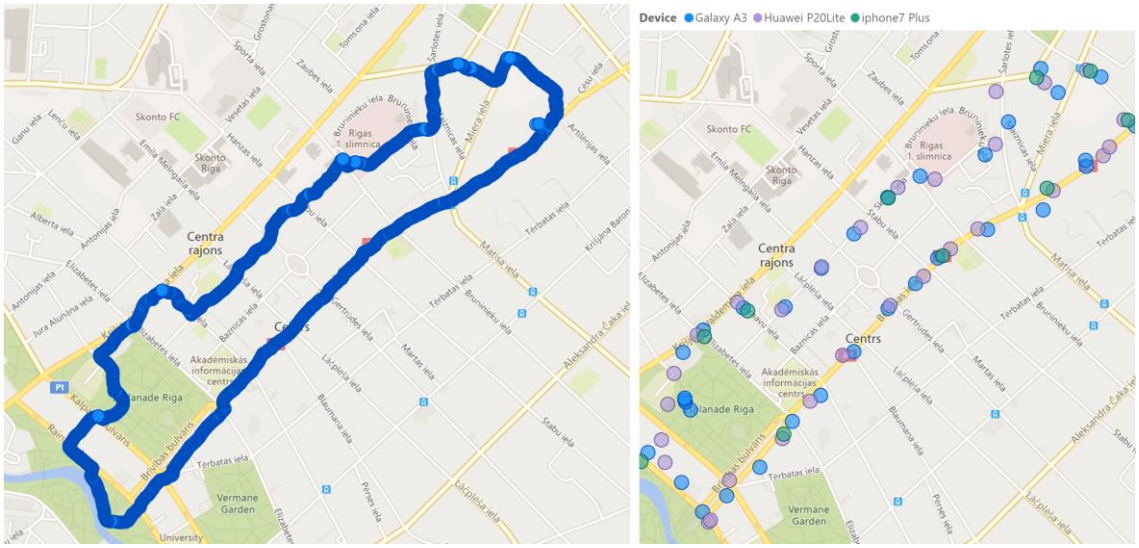


Figure 18. Experiment 1: Validation and test location data on a map.

Experiment 2 presents the results of the Google Maps timeline when the mobile network is disabled. Table 14 presents device configuration in experiment 2.

Table 14. Experiment 2 device configuration.

Experiment Nr.	Location services	Wi-Fi	Wi-Fi Scanning	Mobile network	Bluetooth	Bluetooth Scanning
2.	ON	ON	ON		ON	ON

An important observation was found that it may take up to 30-60 minutes for the phone location data to be loaded on the Google Maps application after being connected to a network. Galaxy A3 performance is similar to the experiment 1 outcome with 28 collected points. iPhone 7 Plus collected 12 data points, compared to 11 data points in experiment 1, and Huawei P20 Lite 1 data point. From the perspective of accuracy, raw data point locations are less precise. The average accuracy of raw data points is 43 meters, and the haversine distance is 44 meters. Compared to experiment 1, accuracy is almost 50% lower when the mobile network is disabled. The single location collected by Huawei P20 Lite has average Google accuracy of 600 meters, indicating that the device gathered the one location data point from a cell tower. Figure 19 exhibits the results of experiment 2.

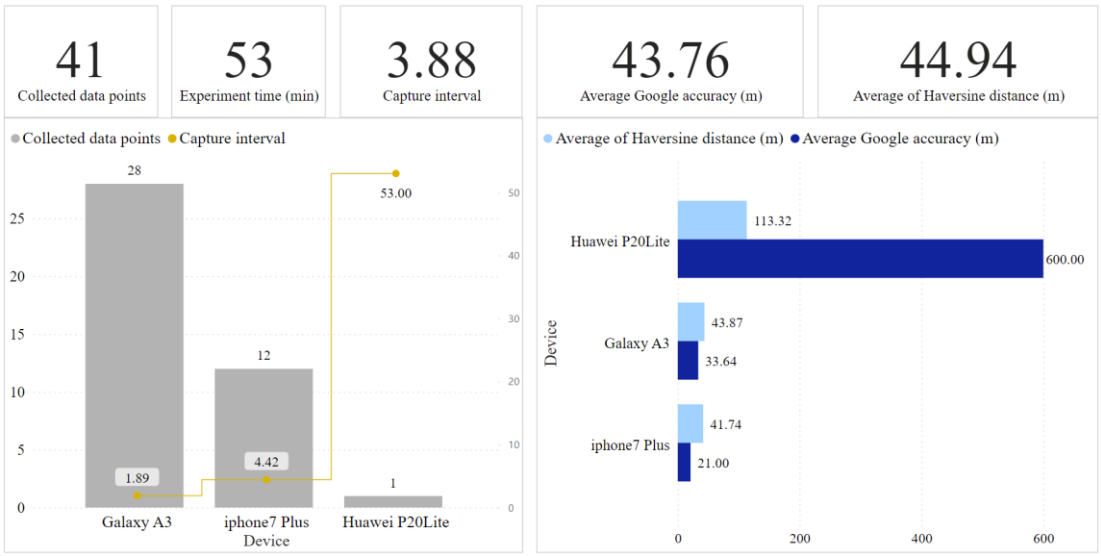
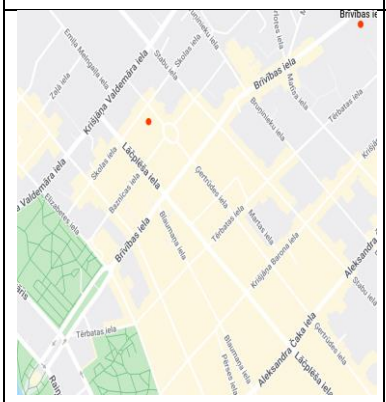
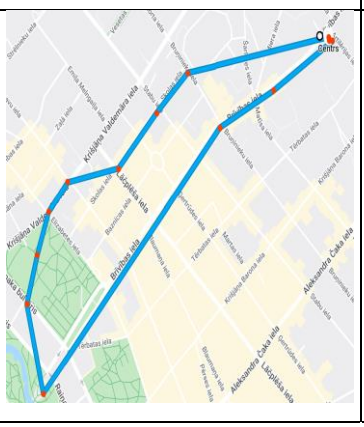
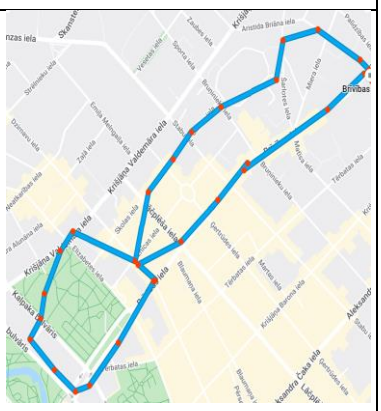


Figure 19. Experiment 2: Raw data point analysis.

The timelines (Table 15) present two paths that are generally precise and a case where the device has not collected a sufficient amount of data to generate a timeline. The timeline on iPhone 7 shows an approximation of the actual route. The timeline on Samsung A3 is more detailed than the iPhone 7 timeline; however, it contains one significant error due to one very inaccurate raw data point. The timeline on Huawei P20 is not presented.

Table 15. Experiment 2: Google Maps timeline.

Huawei P20 Lite	iPhone 7	Samsung A3
		

During the 53 minutes, only one location data point was collected on Huawei P20 Lite. We believe that this may be a one-time error. Therefore experiment 2 iteration two is performed to test this device configuration again. Device Huawei P20 Lite location accuracy was calibrated before the experiment.

In the iteration two thirty-five raw location data points were collected, with a capture interval of 4.71 points per minute. Galaxy A3 collected twenty data points slightly less than in iteration 1. iPhone 7 collected precisely the same number of points - twelve. Huawei P20 Lite collected three raw data points, indicating that this device performs significantly worse than other devices. Similarly to iteration one, Google accuracy and haversine distance are significantly higher than when mobile data is enabled. The average Google accuracy is 172 meters, and the haversine distance is 45 meters. When comparing iteration 1 and iteration 2, the iOS device presents a consistent performance. iPhone 7 average Google accuracy is 27.75 meters, and the average haversine distance is 41.28 meters, compared to 21 and 41.74 meters in the previous iteration. However, android device locations are less accurate. Samsung Galaxy A3 presents average Google accuracy of 244 meters and a haversine distance of 83 meters. Huawei P20 Lite presents Google accuracy of 272 meters and haversine distance of 113 meters. Figure 20 presents the results of the second iteration of the experiment 2.

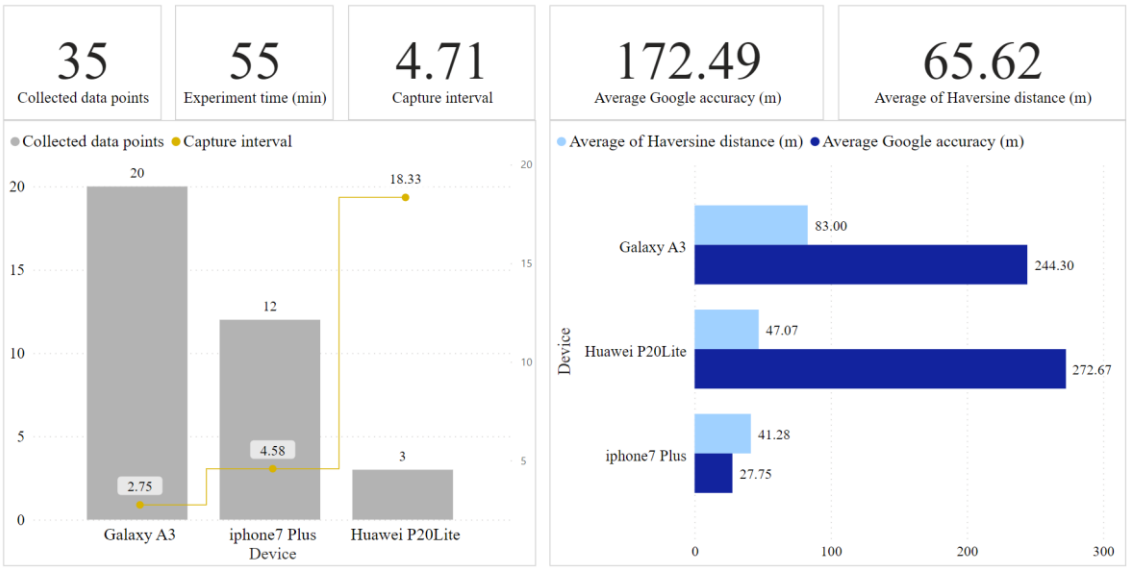
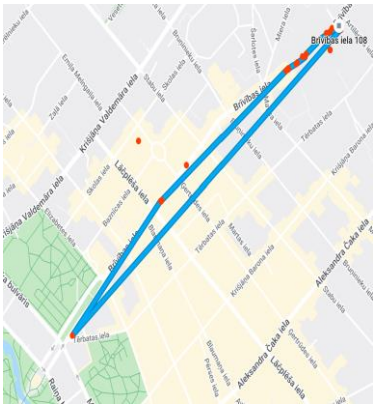
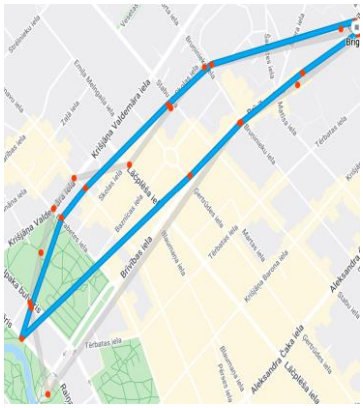
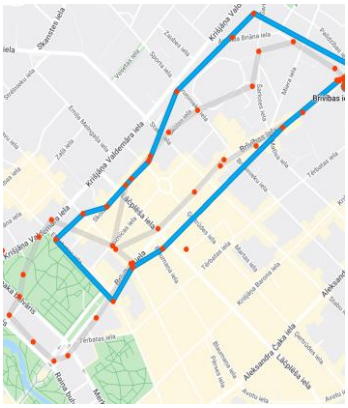


Figure 20. Experiment 2.2: Raw data point analysis.

Manually observing the Google Maps timelines in Table 16 we see three very different paths. Huawei P20 Lite with collected three raw data points presents a very inaccurate route. We could only propose that the device was moving in a specific direction from this screen capture. The iPhone 7 timeline image does not present the major turns in the route. The Samsung A3 timeline route is exact at some parts of the route; however, part of the route is not presented due to missing raw data locations for a specific period.

Table 16. Experiment 2.2: Google Maps timeline.

Huawei P20 Lite	**iPhone 7**	**Samsung A3**
		

The test data points are more randomly collected when compared to experiment 1. There is no pattern, and more points are located outside the real route as exhibited in the right side map of Figure 21.

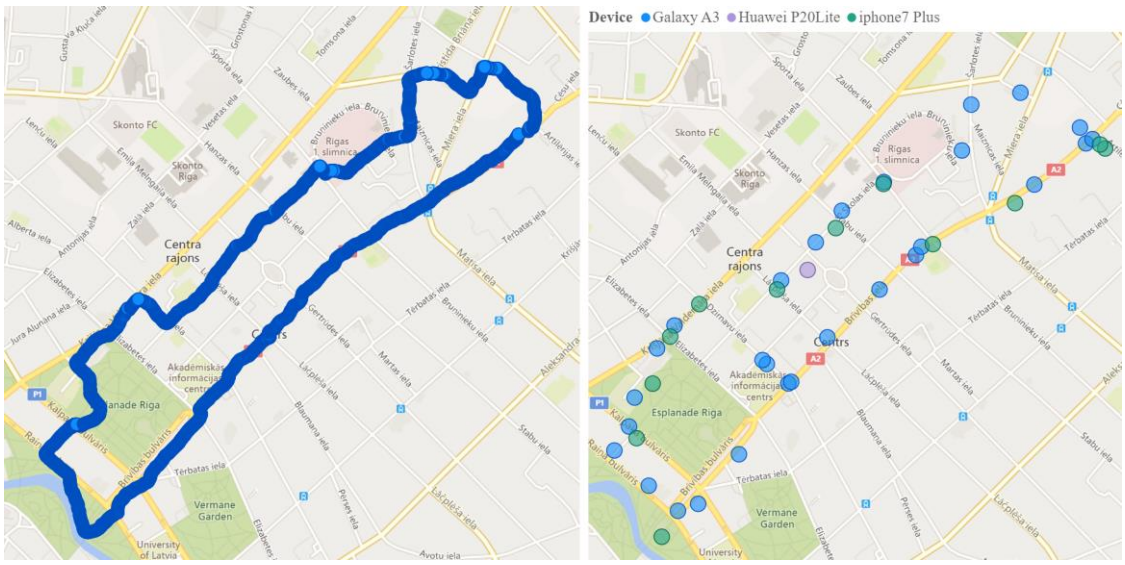


Figure 21. Experiment 2.2: Validation and test location data on a map.

The overall performance of the Google Maps timeline when a device is not connected to the mobile network is sufficient to say that the device was moving in a particular direction; however, the precise route cannot be confirmed. Android and iOS devices perform differently. For an iOS device, there are no significant differences between the experiments with and without enabled mobile data settings. Also, the average Google accuracy and distance from validation locations are similar. These measurements indicate that iOS devices do not use mobile data for location approximating. Android devices perform less consistently, significantly deteriorating the location accuracy when the mobile network is disabled. The average accuracy and distance are 50% worse in both experiment iterations when compared to experiment 1. The applications performance on Android devices is highly affected by mobile network availability.

Experiment 3 is conducted when location services and mobile networks are enabled, disabling Wi-Fi and Bluetooth networks on all devices. We also disabled location enhancing configurations on Android devices during this experiment – Wi-Fi Scanning and Bluetooth Scanning. Wi-Fi and Bluetooth scanning improve the positional accuracy by scanning for Wi-Fi or Bluetooth networks, even if Wi-Fi or Bluetooth are turned off. Table 17 presents the device configuration in experiment 3.

Table 17. Experiment 3 device configuration.

Experiment Nr.	Location services	Wi-Fi	Wi-Fi Scanning	Mobile network	Bluetooth	Bluetooth Scanning
3.	ON			2G/3G/4G		

iPhone collected only 4 location points, compared to 11 or 12 in the previous experiments. Android devices collected more data points than in the previous experiments, with significantly worse accuracy. The average Google accuracy is 493 meters, and the average haversine distance is 96 meters. The result is 94% and 77% worse than in experiment 1. And 78% and 43% worse than in experiment 2. A significant difference is between iOS and android device performance. The average Google accuracy for iPhone is 9.5 and for android is 523 meters. The collected data expose that iOS devices try to collect less but exact data; however, android devices collect many data points with worse precision. Figure 22 presents the results of experiment 3.

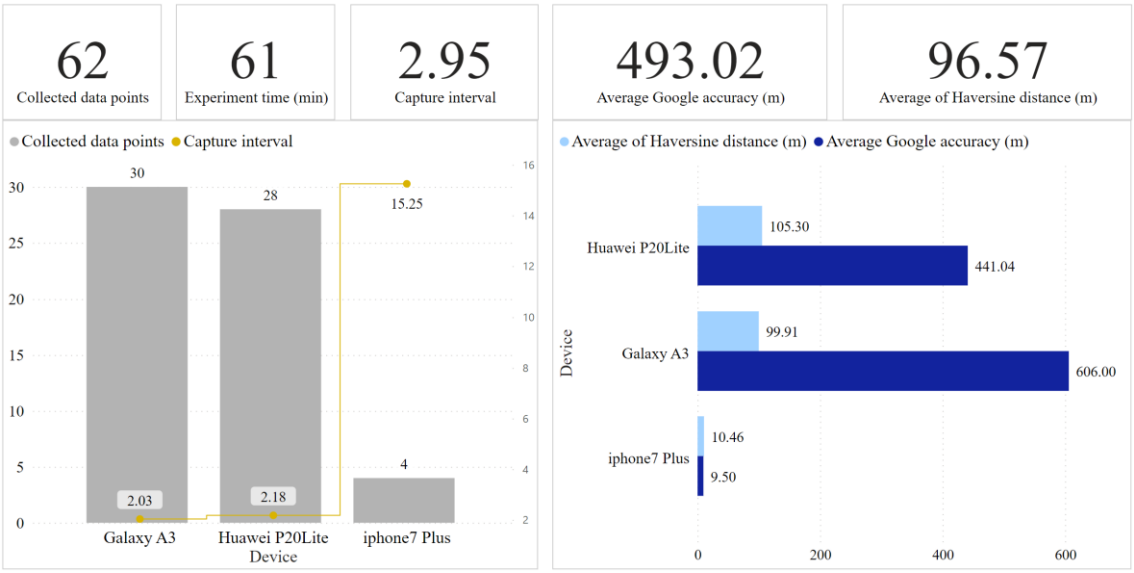
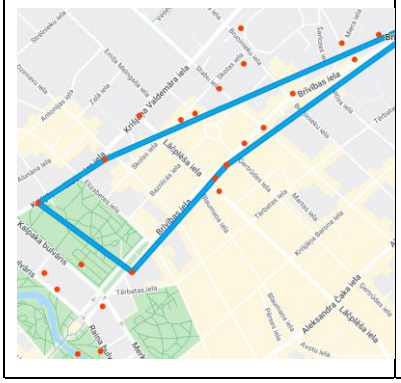
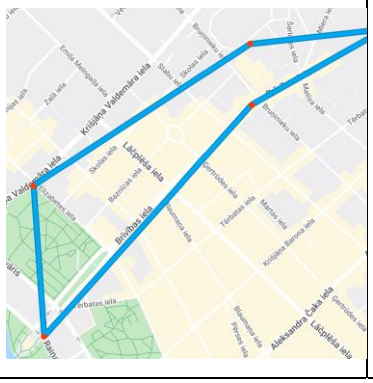
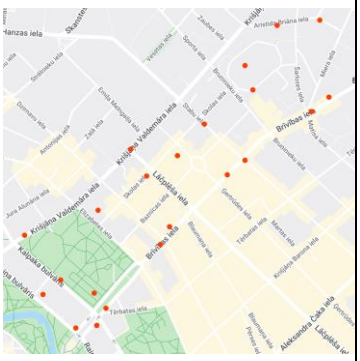


Figure 22. Experiment 3: Raw data point analysis.

We detect a substantial deterioration of the Google Maps timeline performance in experiment 3 as presented in Table 18. Google Maps could present an estimated timeline only for Huawei and iPhone devices. For Huawei P20 Lite, only 6 points from total collected 28 are presented on the actual timeline. For the iPhone device the collected 4 data points make up the timeline and there are no random location points collected during the experiment. Samsung A3 raw location points are with very low accuracy hence Google Maps could not present a route at all.

Table 18. Experiment 3: Google Maps timeline.

Huawei P20 Lite	iPhone 7	Samsung A3
		

Test location data map presents a significant number of raw data points that are not on the route as presented in the left side map of Figure 23. These random data points decrease the timeline's ability to present a route because it increases the complexity of connecting the raw data points logically. The Google Maps timeline cannot distinguish between data points on the route and those that are void.

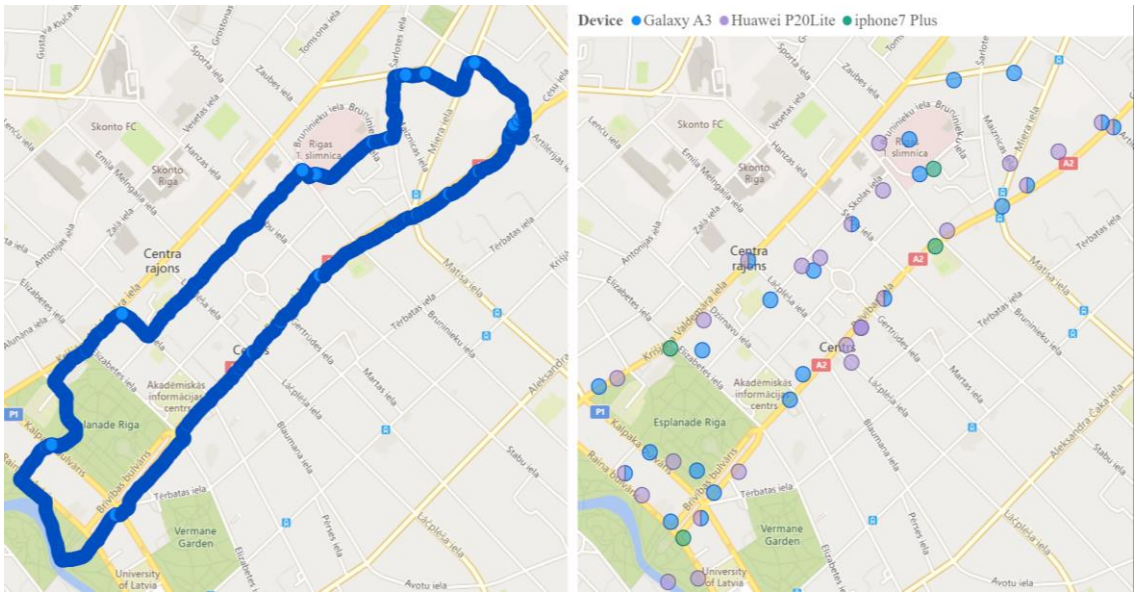


Figure 23. Experiment 3: Validation and test location data on a map.

The Google Maps timeline is performing significantly worse with only location services and mobile network navigation sensors enabled. Many random raw data points are collected that are misleading for Google Maps algorithm to present a precise timeline. There is a significant difference between android and iOS device performance. iOS collects less but highly precise locations. Android collects many data points, with very

low accuracy. The presented timelines are wildly inaccurate and insufficient to say that the device was at a particular location.

In experiment 4, we tested how well the Google Maps timeline can perform when all mobile networks, Wi-Fi, and Bluetooth (including scanning) are disabled and only available enabled setting is the location services. Tested device configuration is presented in Table 19.

Table 19. Experiment 4 device configuration.

Experiment Nr.	Location services	Wi-Fi	Wi-Fi Scanning	Mobile network	Bluetooth	Bluetooth Scanning
4.	ON					

Each device behaved very differently. Samsung A3 tries to locate the device as much as possible with very low accuracy. Huawei located the device two times with a haversine distance from the validation data point – 137 meters. iPhone 7 locates the device four times with very high accuracy. Average Google accuracy is similar to the average haversine distance between test and validation data points. Average sample capture interval is 5.06. It ranges from 2 minutes for Samsung to 29.5 minutes for Huawei. iPhone device collects raw location data on average every 14.75 minutes. Experiment 4 results are exhibited in Figure 24.

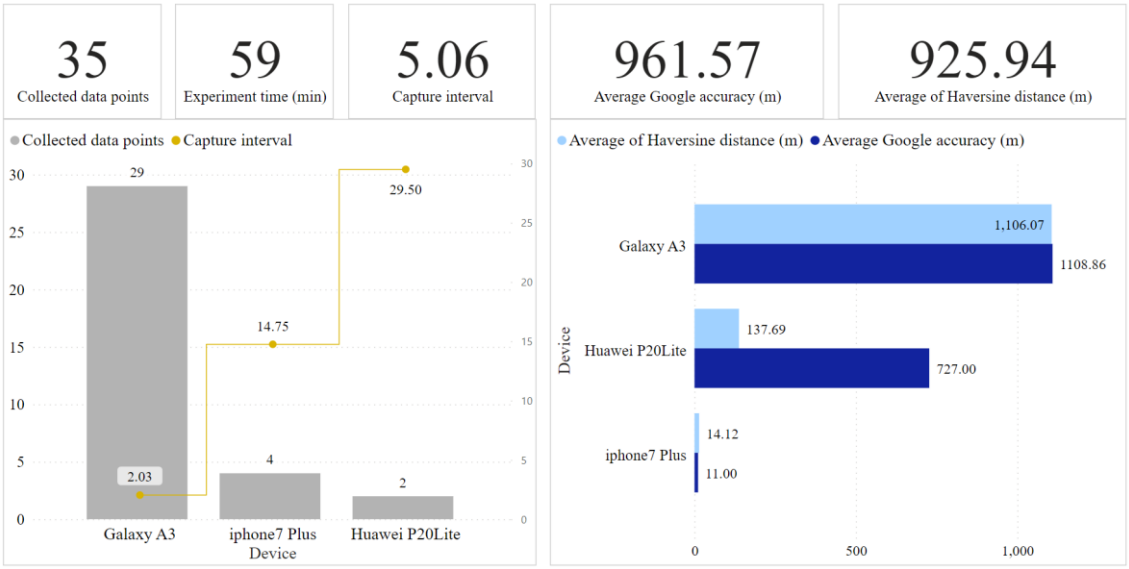
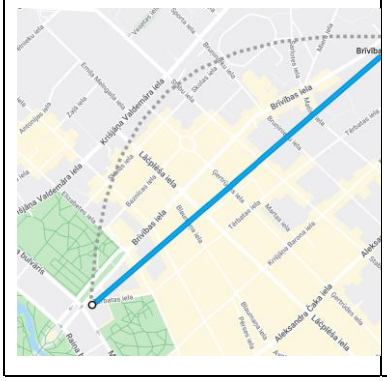
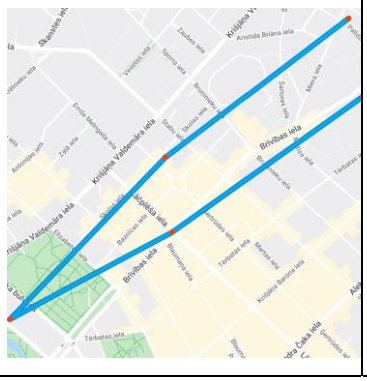


Figure 24. Experiment 4: Raw data point analysis.

With only location services enabled android devices could not present the actual route. Huawei located two raw data points within the experiment time frame, drawing a straight line in the timeline. Samsung A3 located 29 data points, where most of the points are located around the place where the device was disabled from the network. Additionally, for Samsung A3 all data points sources are CELL towers; therefore, average accuracy and distance from validation data points exceed 1000 meters. iPhone collected four data points with very high accuracy. However, it is crucial to outline that a timeline of a path cannot be precisely drawn from 4 raw data points. When only location services are enabled on the device, Google Maps timeline cannot present an accurate timeline as exhibited in Table 20.

Table 20. Experiment 4: Google Maps timeline.

Huawei P20 Lite	**iPhone 7**	**Samsung A3**
		

Each device behaves differently. The most accurate timeline is presented on an iPhone device. Android devices do not present a logical timeline with sufficient and accurate enough data.

Location services were tested in the second experiment iteration to understand the consistency of the results. Huawei and iPhone 7 collected four data points each. iPhone accuracy is as good as in the previous iteration. Huawei collected four points where three of them are located in the same place with accuracy of 35 meters, and haversine distance between test and validation location of 27 meters. The second iteration of experiment 4 results are presented in Figure 25.

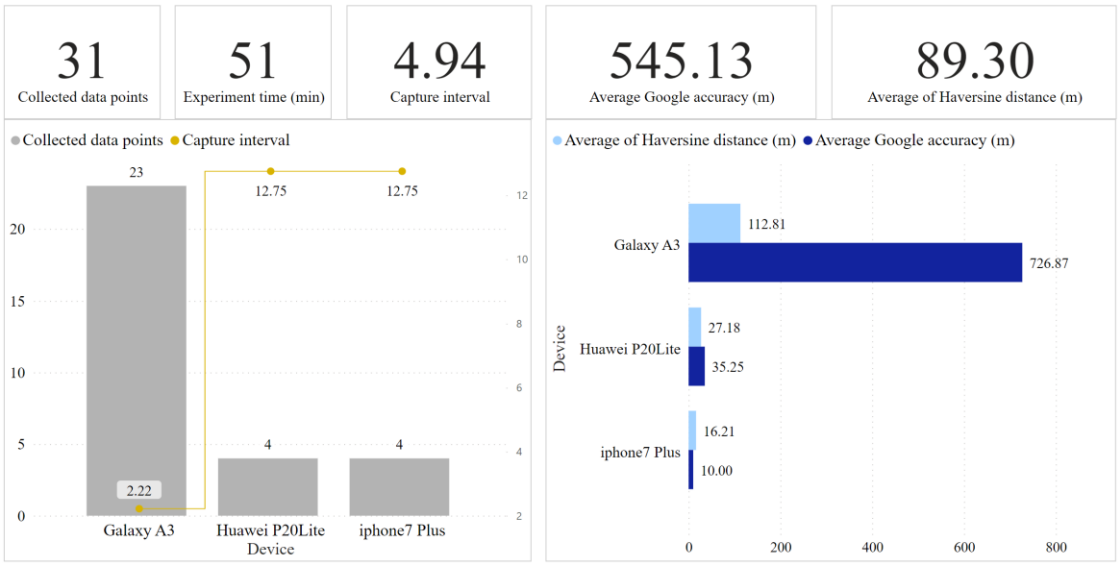
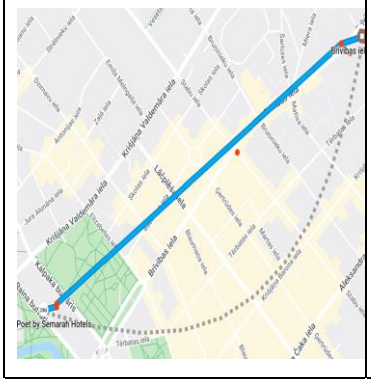
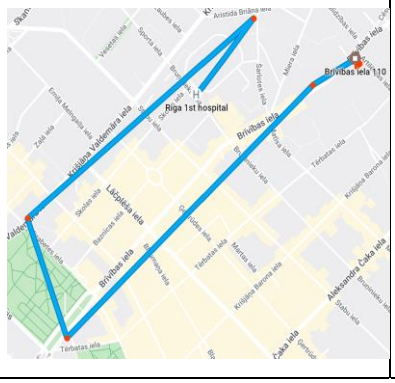
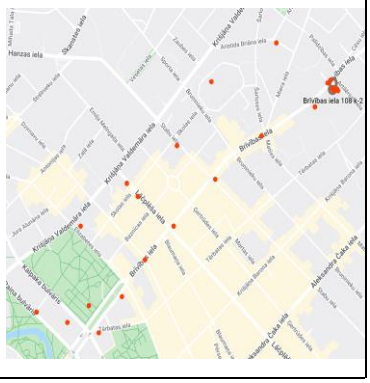


Figure 25. Experiment 4.2: Raw data point analysis.

Exactly the same as in the previous iteration, timeline on Huawei P20 device draw one straight line through two raw location data points, indicating that the device was moving in this direction. The timeline also indicates one visit of "Grand Poet by Samarah Hotels". This hotel is in the path of our route, however it was never visited. In this iteration iPhone 7 timeline very vaguely indicates the actual path. Also, the timeline presents a visit of "Riga 1st Hospital", however it was also never actually visited, only passed by. For Samsung device similarly as in the previous iteration and experiment 3 (Location services +Mobile network) no timeline could be constructed. Even if the raw data points are on the experiment path, the accuracy radius around each raw data point is too big to draw a logical timeline. Generated timelines by the Google Maps application are presented in Table 21.

Table 21. Experiment 4.2: Google Maps timeline.

Huawei P20 Lite	**iPhone 7**	**Samsung A3**
		

4.1.2 Summary

Four different device configurations were tested with their effect on Google Maps timeline performance. In total, 6 experiment iterations were performed.

Raw data capture interval is as crucial as location accuracy because the timeline can only be presented on a map if enough points are collected. When all applicable device settings are enabled, the capture interval is 2.63—indicating that a device will capture a new location point every 2.63 minutes. By disabling the settings one by one, the capture interval increases as displayed in Figure 26.

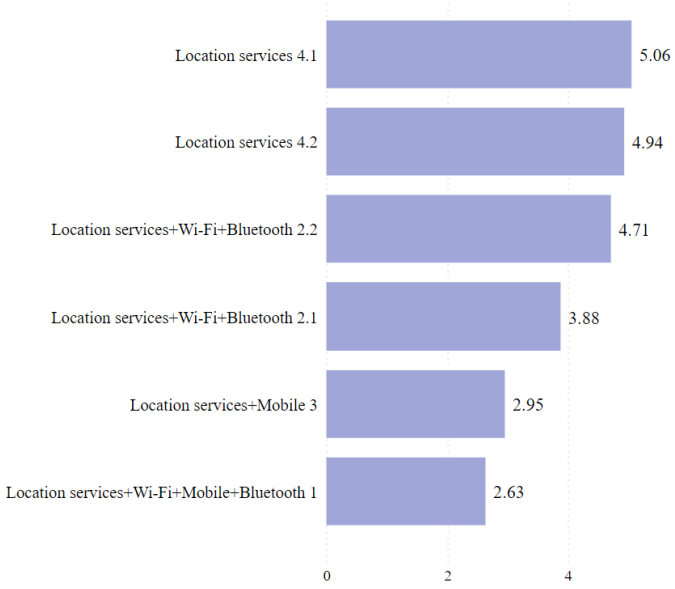


Figure 26. Sample interval summary.

Figure 27 presents the Google accuracy against the haversine distance. There is a direct correlation between these variables. When Google accuracy is large, the haversine distance also increases.

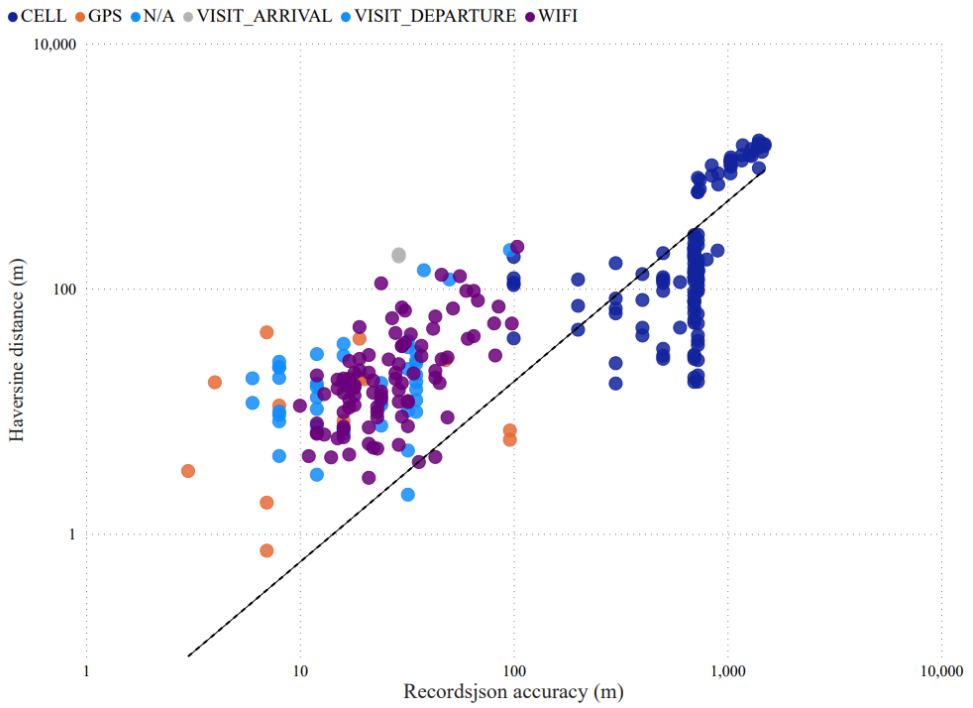


Figure 27. Raw location data point overview.

The source of location "CELL" holds the worst performance. The two clusters of cell tower raw locations are two different device configurations. The highest cluster presents data points when only location services are enabled with a haversine distance and Google accuracy of around 1 000 meters. However, the dark blue cluster below the trendline presents data when location services and mobile networks are enabled. The majority of these raw data points hold an accuracy of 600 meters. The actual distance to the ground truth location for these points is around 100 meters indicating that Google Maps is very pessimistic when retrieving data from cell towers, even if a mobile network is enabled. Wi-Fi source raw location points hold accuracy between 10 to 104 meters and haversine distance between 3 to 220 meters. There are very few GPS source data points with varying accuracy and distance. The most accurate retrieved point is with an accuracy of 7 meters and a distance to validation location of 1 meter.

N/A data points are from the iOS device, and they hold the highest median accuracy of 16 meters and the same median haversine distance of 16 meters. 50% of all accuracy values lay within the 12-to-32-meter range, and haversine distance values are within the 10-to-23-meter range. iOS device presents an excellent approximation of location detection with a stable certainty. There is significantly less noise in the data with only exact locations. There are three location sources in Android devices: CELL, Wi-Fi, and

GPS. Considering the average accuracy and the haversine distance, the worst performance is for the cell tower source. The median Google accuracy for cell towers is 699 meters, and the median haversine distance to the validation location is 136 meters. 50% of all CELL source accuracy values are between 699 to 727 meters and haversine distance between 76 to 275 meters. GPS source performance for both median accuracy and haversine distance is 17 meters. GPS source holds the lowest inter quartile range of 12 and 13 meters. Wi-Fi source holds the same haversine distance as GPS source, however median Google accuracy is 7 meter higher. Google accuracy and haversine distance by the source of the location are presented in Table 22.

Table 22. Experiment result overview on the location source level.

Source of location	**Median of accuracy (m)**	**1st quartile of accuracy (m)**	**3rd quartile of accuracy (m)**	**IQR (m)**
N/A				
Google accuracy (m)	16	12	32	20
Haversine distance (m)	16	10	23	13
GPS				
Google accuracy (m)	17	7	20	12
Haversine distance (m)	17	6	19	12
Wi-Fi				
Google accuracy (m)	17	10	34	24
Haversine distance (m)	24	17	37	20
Cell				
Google accuracy (m)	136	76	275	198
Haversine distance (m)	699	699	727	28

The worst performing raw location data are collected when only location services are enabled. Google Maps suggests accuracy of 961 meters. The calculated haversine distance between these locations and the validation dataset is 956 meters. Similar results between Google accuracy and haversine distance demonstrate the correctness of the accuracy assumption. When a mobile network is enabled majority of the data points are gathered from cell towers with significantly higher accuracy. Google's accuracy is more pessimistic than the actual precision of the raw locations. The average haversine distance is around 50% smaller than Google's accuracy. Mobile network data significantly

improves the device's ability to gather precise locations. A combination of location services, Wi-Fi, and Bluetooth collects raw data with a haversine distance of 55 meters from the validation dataset and Google accuracy of 109 meters. With these device configurations, most data points have a Wi-Fi source, with some points located from cell towers. The best performance is when all navigation sensors are enabled. Haversine distance of 22 meters and Google accuracy of 29 meters present a very close approximation of the actual location. The average Google accuracy and haversine distance are presented in Figure 28.

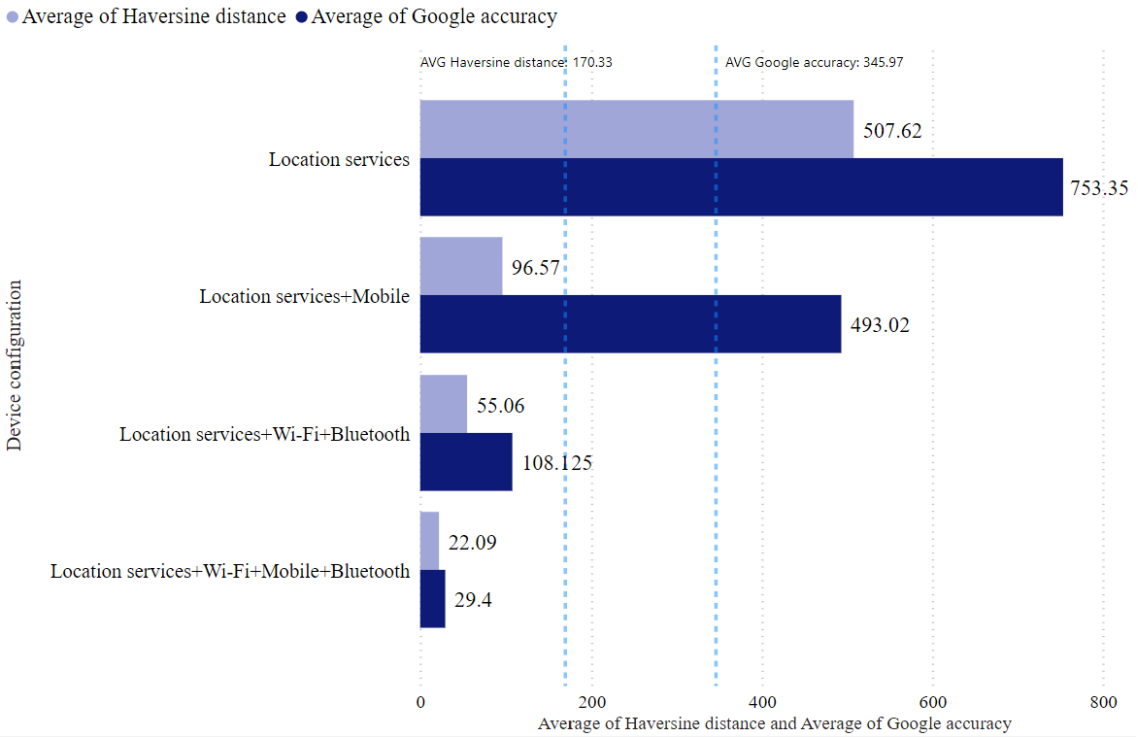


Figure 28. Experiment result overview on the device configuration level.

Several additional observations are made during the experiments. Firstly, if the device is not connected to a mobile network or Wi-Fi, it takes above 30 minutes for the device and Google Account to synchronize the location data. Secondly, iOS and Android devices behave substantially differently in various settings due to different device features.

4.2 Place visit identification accuracy

In the town of Balvi, ten different locations were visited. Appendix 1 presents detailed detected visited location summary. Figure 29 on the left side map presents the visited locations numbered respectively and on the right side are the raw data location points

collected by the mobile devices. The visited locations were positioned in the town (4;5;6;7) and in rural areas (1;2;3;8;9;10).

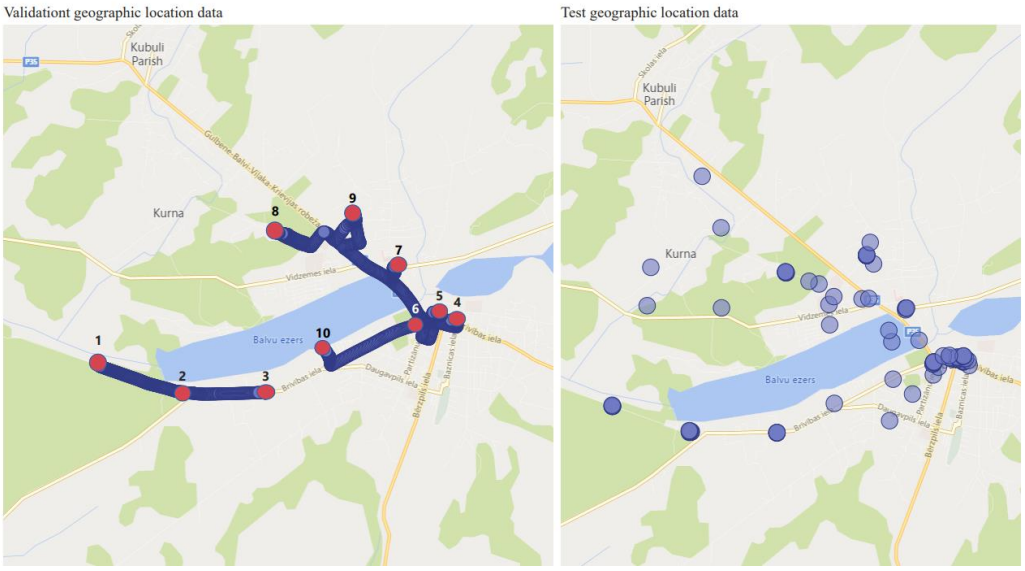


Figure 29. Visited places on a map, Balvi.

The dwell time in each location was no less than 10 minutes. The results show that only locations positioned in the town and with a semantic meaning were presented in the timeline as visits. The devices detected none of the locations in the countryside with or without an address. The stops not presented in the timeline were only shown as if the device was moving through them without stopping. The Google Maps timeline detected 50% of all visits in the town area (Figure 30).

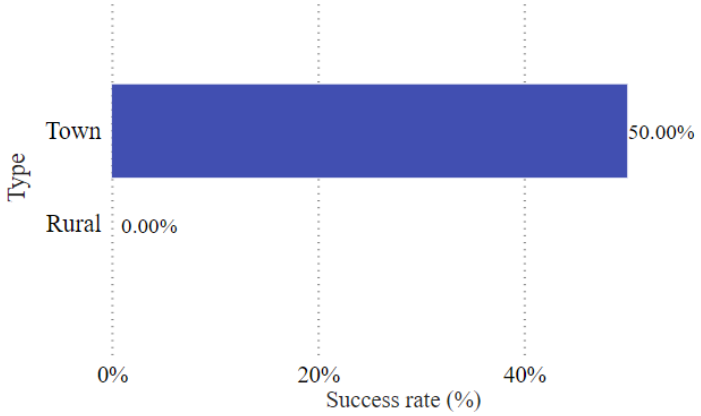


Figure 30. Visit success rate (%) by type, Balvi.

Analyzing the rural area places not detected by the devices, we observe that an insufficient number of raw location data points does not seem to be the main reason these visits are

not detected. Table 23 presents two visits with respective collected raw data points. First visit was not detected and second visit was detected in the timeline.

Table 23. Raw data points in visits 2;7, Balvi.

Visit 2, rural area (not detected)	
Visit 7, town area (detected)	

Visit number two does not have an address or a semantic meaning, only physical location such as latitude and longitude. During the 10 minutes, 13 raw data points are collected. iPhone also presents visit arrival at 10:30:47 and departure at 10:40:18, indicating that there has been a 10-minute dwell time. Visit number seven is in the town center, containing 12 raw location data points. Visit two was not detected by any mobile device, on the contrary, visit seven was detected by two devices (iPhone and Samsung). The only difference between the two locations is that number two is in a rural area without semantic meaning, and number seven is in a town with a semantic meaning. Therefore, leading to a conclusion that the number of raw data points is not the main factor influencing the accuracy of place visit identification. We conclude that the Google Maps timeline does not semantically identify rural area visits. However, the Google Maps application collects

enough data to confirm a visit from the raw location data manually, even if it is not presented in the Google Maps timeline.

For all identified visits, the average visit confidence is 82.6, and the average location confidence is 45. Balvi lake is an identified false positive visit by an iPhone device. This location may be identified because it is between several other visited places. All other identified visits are correct. These visits have varying semantic location confidence, within a minimum of 16 and a maximum of 82. All identified visits and respective location and visit confidence values are presented in Figure 31.

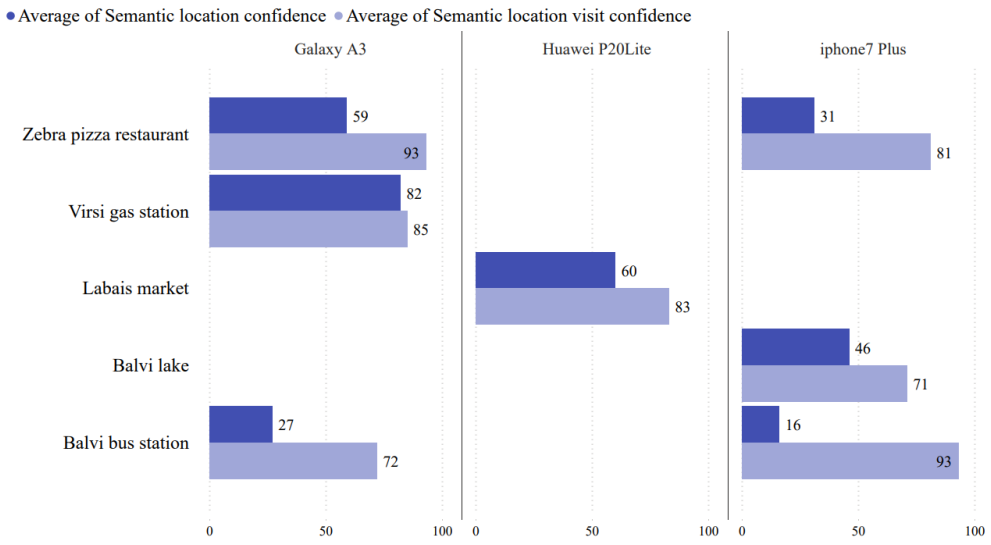


Figure 31. Identified visits by devices, Balvi.

In Riga, ten places were visited. Contrary to the previous iteration, we tested the visit type– inside or outside, not the type of the location. Visited places are presented in Figure 32.

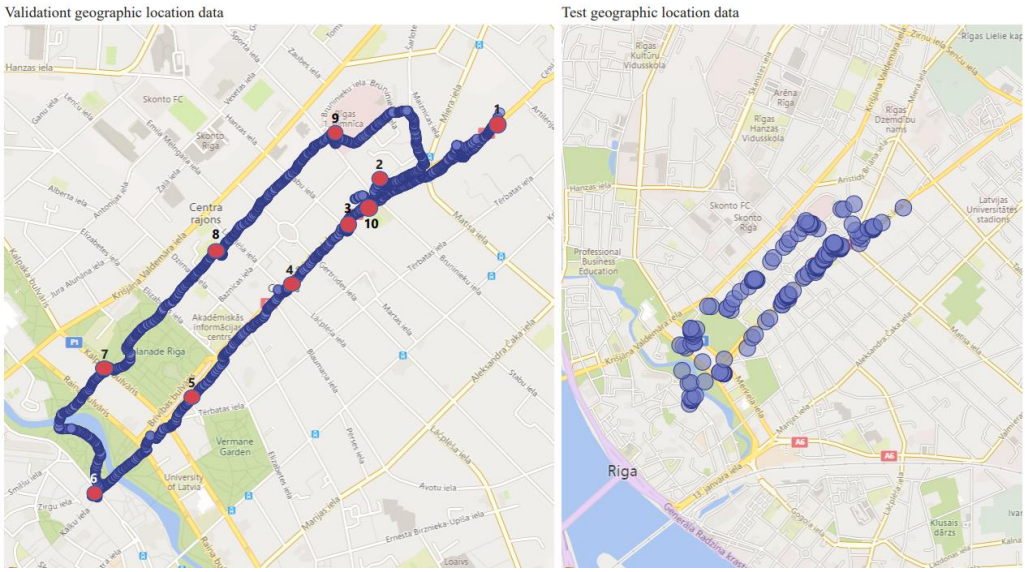


Figure 32. Visited places on a map, Riga.

The success rate for the detected visits by the Google Maps timeline is significantly higher in Riga. 42% of all visits were presented in the timeline, compared to 20% in Balvi. Visits, where we went inside the building hold a considerably higher probability of being identified by the Google Maps timeline. The inside visit success rate is 75% compared to 19% of outside visits. Visit success rate is displayed in Figure 33.

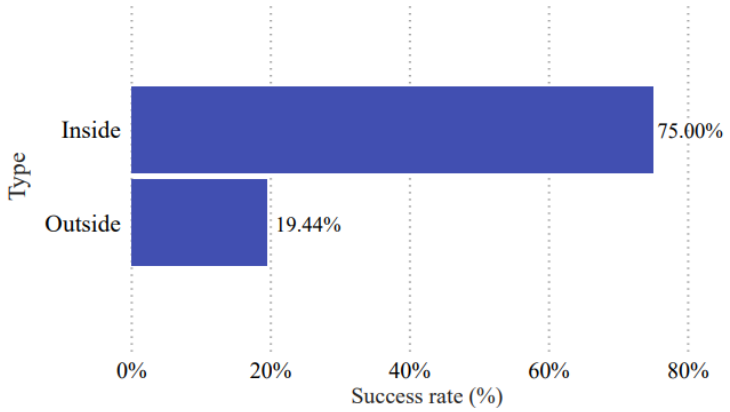


Figure 33. Visit success rate (%) by type, Riga.

Wi-fi access points are a significant source for a place to be identified as a visit, and the majority of Wi-fi access points are connected only when the device is within a specific range of the access point. The access point range is one of the reasons why we have observed that the visits inside the building hold a higher success rate. Android devices detected 100% of inside visits, supporting the idea that Google Maps timeline visited places are accurate and precise when a visit is inside a building, and the visited location

holds a semantic meaning. iOS device performance is not the same, and the Google Maps timeline application did not detect any visits precisely.

Android devices equally identified visits 4;5;6;9;10. Additionally, the Samsung device identified Matisa street as a visit near the Narvesen shop that was the first visit. We observe that the average location confidence is significantly lower for visits that are false positive: “Lauvas nams restaurant”, “Pharmacy Saules” or “St. Alexander Nevsky Church”. iPhone identified locations that are near the actual visits. St. Alexander Nevsky Church is across the street from Drogas shop, Lauvas nams restaurant is also across the street from Daile theater. Bastejkalna park and „Ausmeņa“ kebab are false positives near the actual visit locations. All identified visits are displayed in Figure 34.

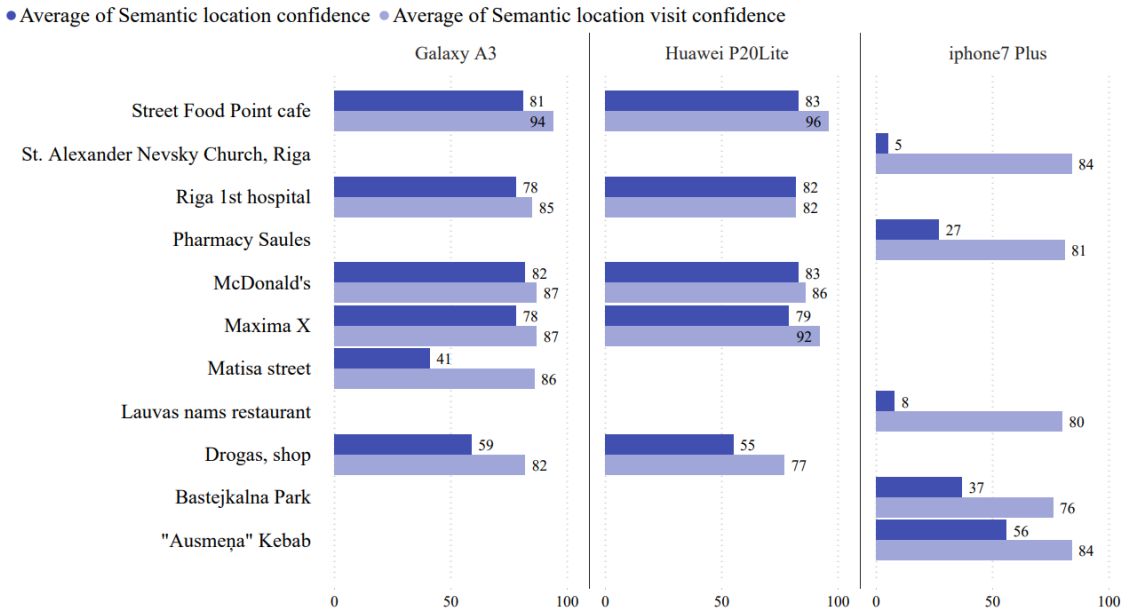
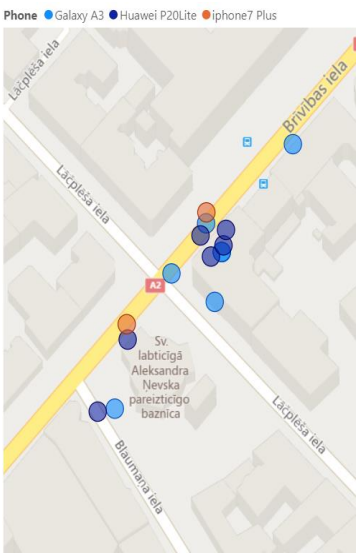
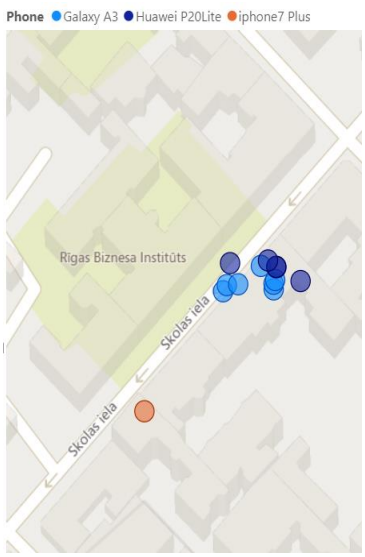
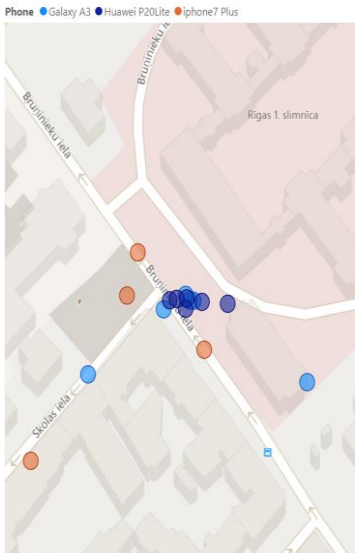


Figure 34. Identified visits by devices, Riga.

Table 24 presents three visits and collected raw data points during the visit time.

Table 24. Raw data points in visits 4;8;9, Riga.

Visit 4 (semantic)	Visit 8 (address)	Visit 9 (semantic)
		

Visit 4 is in the shop “Drogas”. There are 12 raw data points collected during the ten minutes. The Google Maps timeline detected this inside shop visit correctly from devices Samsung and Huawei. The iPhone device collected only two raw data points leading to a false-positive visit to a church near the shop. Visit 8 of a location with no semantic meaning was not detected even if 13 raw data points were collected during the 10 minutes. iPhone detected a false-positive visit to a nearby food shop, “Ausmena” kebab. Visit 9 to Riga's 1st hospital was detected by both Android devices. The raw data map clearly displays a cluster of points next to the hospital. The iPhone device collected four raw data points in the hospital region and did not identify a visit. All three visit raw data point clusters indicate a stop, however, only stops with a semantic meaning are detected by the Google Maps timeline.

Locations with a semantic meaning hold the highest likelihood of being detected and presented in the Google Maps timeline. A location without a semantic meaning was detected only once. Furthermore, no physical-only locations were detected during the experiments. Android devices collect raw data points with lower capture intervals; hence Huawei has detected 50%, Samsung 70%, but iPhone device detected only 30% of the semantic locations. Visit success rate by mobile device and type of location is displayed in Figure 35.

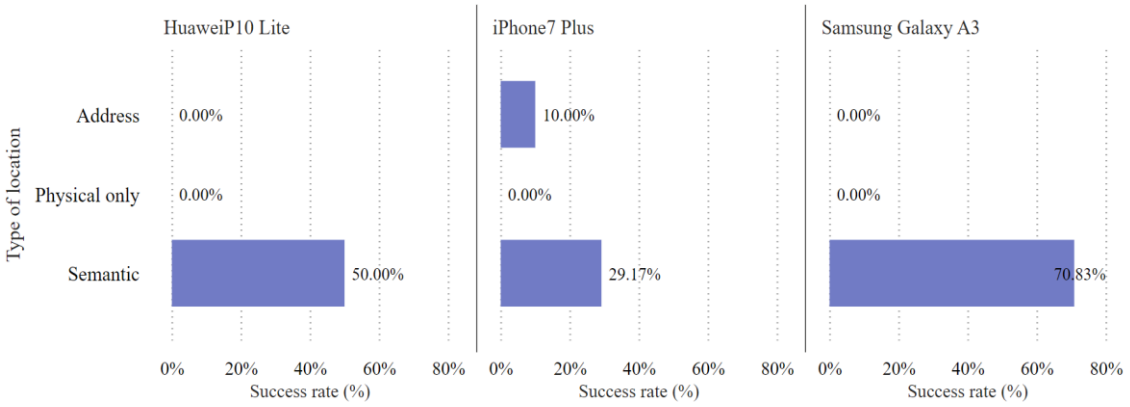


Figure 35. Visit success rate by the type of location and device.

Several factors must be highlighted when summarizing the Google Maps timeline performance regarding place visit identification. Location type is an essential factor for a visit to be identified in the application. Visits in rural areas will not be presented in the Google Maps timeline, even if the dwell time is above 10 minutes. Indoor visits have a significantly higher likelihood of being detected than locations in outdoor settings. Wi-Fi access point scanning is the primary method of identifying location visits. Raw location data points after extraction and transformation can indicate that a place has been visited, even if it is not presented in the Google Maps timeline. Android devices perform differently from iOS devices. iOS device performance is significantly worse and cannot present actually visited locations. When the algorithm is unsure about the visit, location confidence genuinely presents a low location confidence value. The value is low for the visits that are mainly identified wrongly. Location confidence and raw data points are useful indicators when recognizing if a visit has happened.

5 Discussion

The discussion chapter presents a forensically sound data acquisition protocol from the Google Maps timeline application using a mobile device. We capture and consider the accuracy of each extraction layer as discovered in the experiments and outline specific points of interest where inaccuracies may arise. Additionally, we comment on the use cases and limitations associated with the experiment results.

NIST presents a mobile device tool acquisition pyramid with five layers: manual extraction, logical extraction, JTAG, Chip-Off, and Micro Read [65]. Each technique requires a different level of expertise and tools. An explicit comparison of logical acquisition techniques combined with vendors and physical acquisition techniques is presented by S. C. Sathe and N. M. Dongre in [66]. We propose a four-layered data acquisition approach (Figure 36) for each location data type available in the Google Maps timeline depending on the access level to the Google account, mobile device condition, and case unique characteristics.

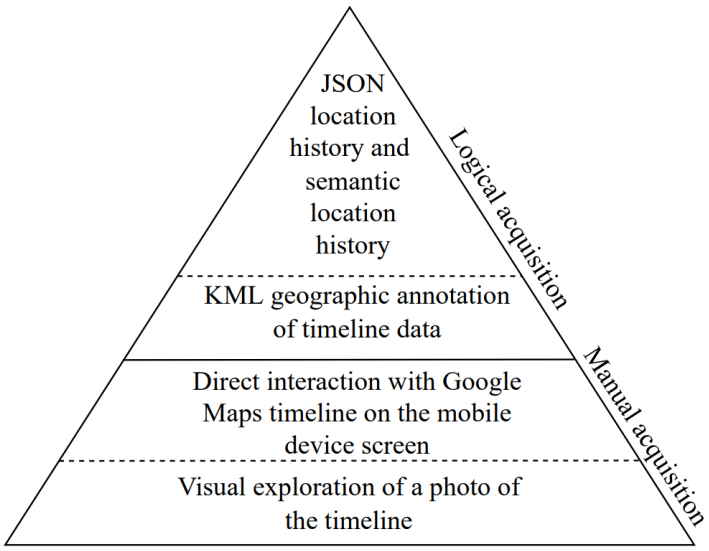


Figure 36. Data acquisition methods and respective data types.

As mentioned by Casey and Rose, forensic examiners must hold a solid knowledge of the technology that edits, arranges, interprets, and displays the underlying forensic data [33]. Google Maps timeline arranges and interprets the raw location data to be understandable

for the end-user in the timeline. The vaguest and most unreliable data level is a visual inspection of the Google Maps timeline in the application or through screen captures of the timeline. Forensic examiners must understand the limitations and algorithms behind the displayed data while interacting directly with the app or visually examining a photo of the Google Maps timeline. KML raw data exported from the application is a geographic interpretation of the timeline as seen in the application. Most detailed and unprocessed location data is retrieved as JSON files from the Google Takeout service.

The manual data acquisition layer is linked with a direct inspection of the timeline in the app and a visual inspection of the timeline images. Logical acquisition layer is linked with KML and raw JSON file acquisition and analysis.

5.1 Manual data acquisition

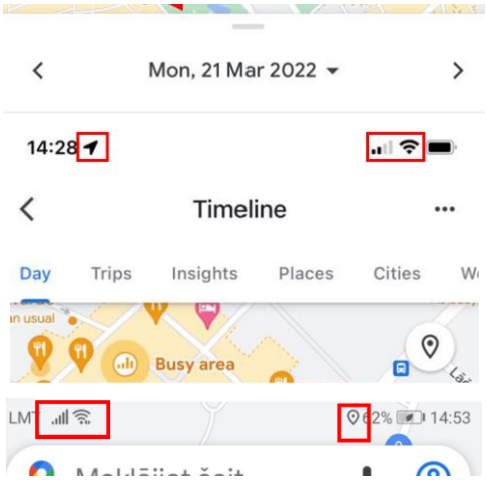
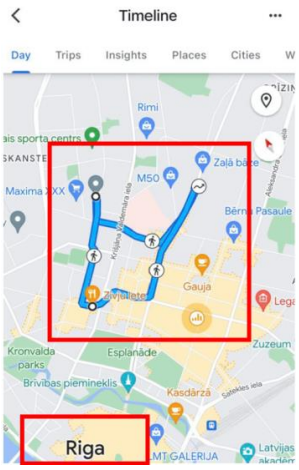
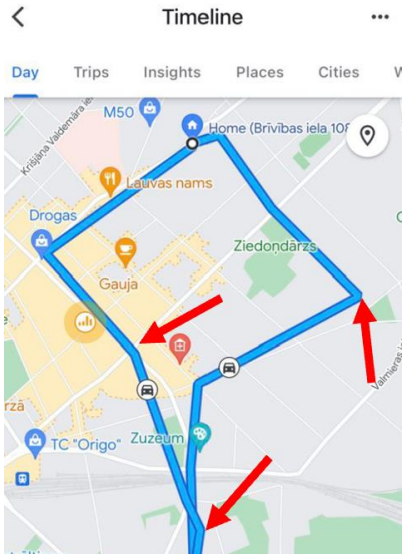
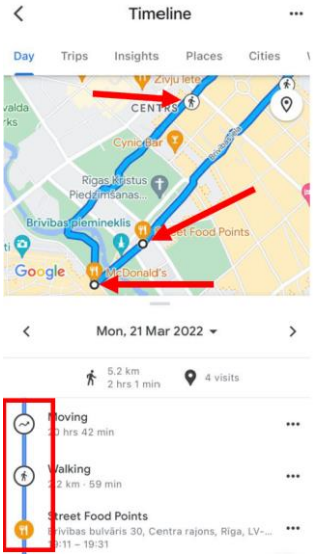
NIST defines manual data acquisition as employing the user interface to retrieve the information directly on the mobile device screen [65]. Google Maps timeline can be viewed on the phone using the screen captures of the map or directly by interacting with the app that is running on the device.

Visual exploration of a screen capture of the timeline

1. **Data interrogation**

The Google Maps timeline screen capture can be manually observed on the individual's mobile device or after acquiring an image of the device's memory. Additional software or hardware is not needed to analyze the presented data. Before looking into the timeline, observe the mobile device's status bar if it is available on the screen capture. Digital investigators must understand what navigation sensors were enabled when the screen capture was made. First, observe how precise are the lines snapped on the roads. The more precise the lines, the more frequently the raw data points have been captured. Second, observe if there are any location visits presented in the timeline by looking at the different icons on the route. Each visit address can be separately viewed in other mapping application and the nearby locations can be detected and analyzed. Third, analyze the overall surroundings, whether the timeline is in a rural area or a city. Depending on the environment, the accuracy of the raw data points and visits will vary. Table 25 exhibits the proposed data interrogation process for visual exploration of a capture of the timeline.

Table 25. Proposed data interrogation process for exploration of a screen capture.

<table>
<tr><td>1. Acquire the screen capture</td><td>2. Create a backup for the screen capture</td></tr>
<tr><td>3. Note enabled navigation sensors (Wi-Fi, Mobile network, Location services) and date of the timeline
 
</td><td>4. Note the environment of the timeline – this affects the accuracy of the route and visit detection
 
</td></tr>
<tr><td>5. Observe the lines snapped on the roads and intersection points between the lines – these are the locations of the raw data points
 
</td><td>6. Observe the icons on the route – they indicate a visit or type of movement and using a different map view these locations
 
</td></tr>
</table>

2. Accuracy

If all navigation sensors are enabled (location services, mobile network, and Wi-fi), then the probability that the device was moving in the given direction is very high. The direct lines between the raw data points present the possible route; however, they should not be used in the analysis. Observe the data points in intersections between the lines because

these points present where the raw data points are actually placed. The visit accuracy depends on the type of the location and the environmental surroundings. There is a high likelihood that most of the visited locations will not be presented in the rural areas. It will be shown that the device moved through the visited location without stopping. In the city environment, locations where the mobile device was inside and that hold a semantic meaning hold a 75% probability of being detected if the visit time is above 10 minutes. Visits of addresses or physical locations are more ambiguous and may not be presented correctly.

3. **Use cases**

This acquisition level is practical only when screen captures of the timeline are available. These can be received or created screen captures discovered on the mobile device or any other storage account. No access to the Google user account is needed, hence we are not required to acquire legal access to the cloud account. The data examination is quick and data manipulation or improper handling issues are more unlikely then for other extraction layers. The locations are only approximate without any accuracy indications. We propose that the data from this layer can only be used to support other evidence rather than being as a distinct finding used as primary evidence.

4. **Limitations**

Google Maps timeline provides the end-user with a summarized data review of the visits and modes of travel. Any data modification or deletion in the Google Maps timeline cannot be distinguished from an actual event on the screen capture. The data attribution is impossible because we cannot confirm who was logged in to the Google account or used the device from a screen capture. Data validation is complicated and close to impossible.

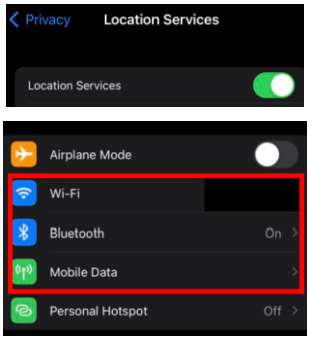
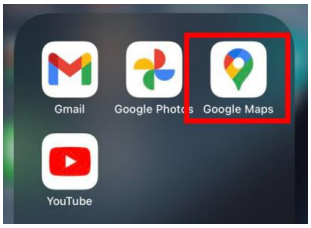
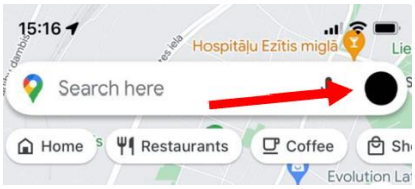
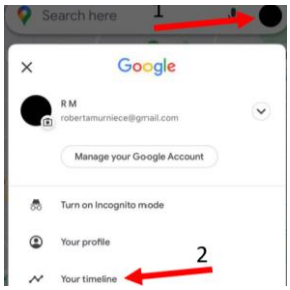
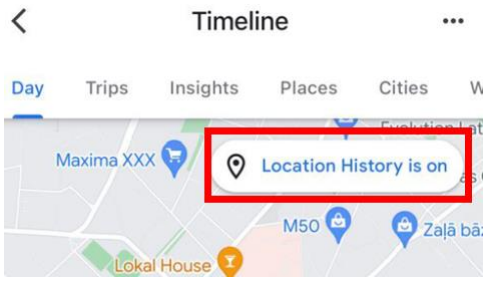
Direct interaction with the Google Maps timeline

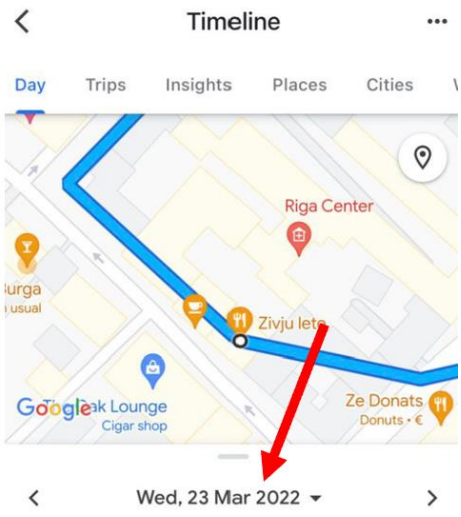
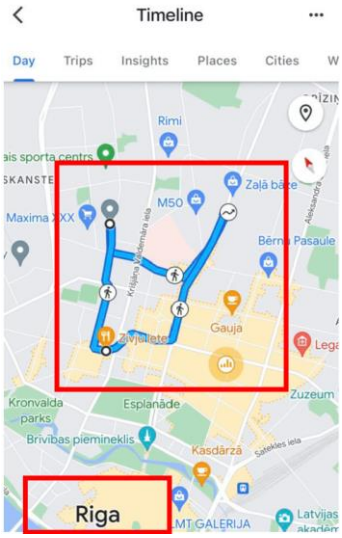
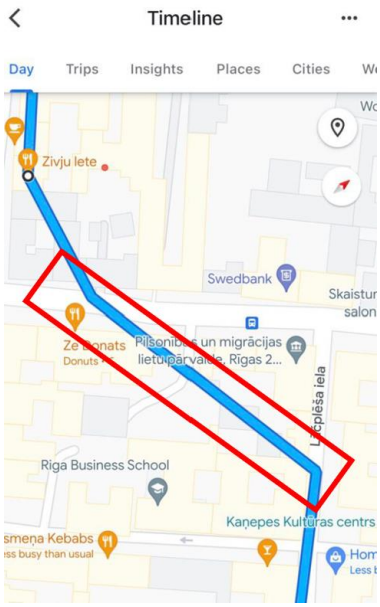
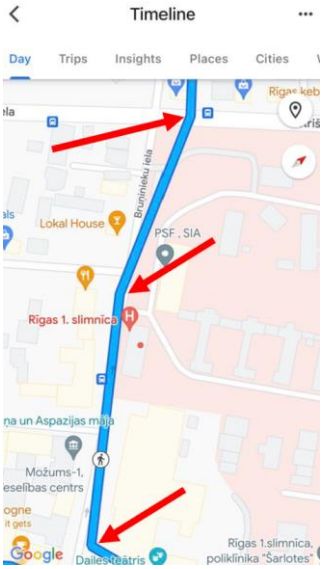
1. **Data interrogation**

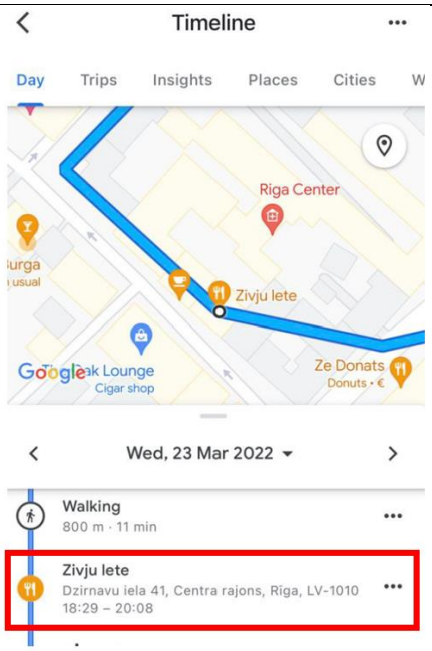
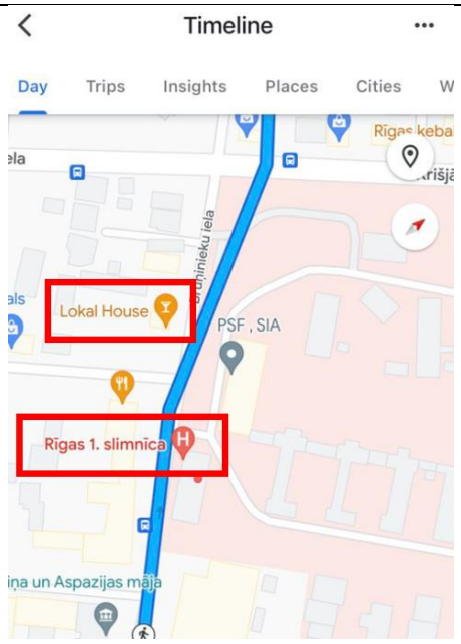
The data interrogation must be documented in a video to present a valid chain of custody process. The video presents the data as it was at the moment of analysis. Google Maps is an application where data is stored in the cloud, hence data modification can be done remotely. The mobile device with a logged-in Google account should not be moved not

to alter the existing timeline. Before interrogating the timeline, it is crucial to document what navigation sensors are turned on the mobile device because it affects the accuracy of the timeline. After opening the Google Maps application, observe what account has been logged in. Afterward, select the day of interest. Observe each event in the timeline separately – a visit or a travel route. First, review where the possible raw location data points are collected. These are points in between the line intersections of the route. Discard the line of the route as it is not valid and accurate. Second, detect the time of each travel or route. Combining raw data with timestamps may determine the device's approximate time at a particular raw data point. Third, observe the environment (rural or city) and visited location types (semantic, address, or physical). These variables directly affect the accuracy of the timeline. Direct interaction with the timeline data interrogation process is exhibited in Table 26.

Table 26. Proposed data interrogation process for direct interaction with the timeline.

<table>
<tr>
<td>1. If possible, do not move the mobile device
2. Turn on the video recorder, record time, name of the investigator and start recording the timeline interrogation</td>
<td>3. Note enabled navigation sensors
 
</td>
</tr>
<tr>
<td>4. Open Google Maps application
 
</td>
<td>5. Record the account that is logged in to the Google Maps application
 
</td>
</tr>
<tr>
<td>6. Open Google Maps timeline
 
</td>
<td>7. Detect location history settings
 
</td>
</tr>
</table>

8. Select the day of the interest 9. Zoom-in the timeline and record the timeline throughout the route 	10. Note the environment of the timeline – this affects the accuracy of the route and visit detection 
11. Observe the lines snapped on the roads – how often the raw data has been captured 	12. Observe and expand the intersection between lines that have different angles – these are the locations of the raw data points 
13. Observe visited locations – duration, type (rural/city or inside/outs), nearby locations	14. Observe the locations that are nearby the route – some visits may not be detected in the timeline

	
15. Save the recording and turn off the video recorder	

2. Accuracy

Similar to the previous extraction layer, the accuracy depends on the enabled navigation sensors and external environment. If Wi-Fi, location services and mobile network are enabled then the raw location data points are collected every 2.63 minutes. This capture interval is sufficient to describe an approximate route. We propose that there is a very high to 100% probability that the device was within 30 meters from the route presented in the Google Maps timeline. We cannot observe if a visit may be missing in this extraction layer because we still cannot see the raw data points. As we observed in the experiments, even if a visit is not detected, there may be clusters of raw data points that should be manually observed indicating a visit, for example, in a rural area. Mainly semantic locations are detected as visits; therefore, for detailed analysis, the list of visits may be inaccurate, especially if visits are in only address or physical locations.

3. Use cases

A manual data acquisition technique can be used to retrieve the information instantly and in cases of non-cooperative circumstances. These include when the owner of the Google account is not providing their log-in credentials. At the same time, the account is logged in the application, and manual data viewing is feasible on the mobile device. Additionally,

this extraction technique can be used to retrieve the data as fast as possible and to support other evidence. The manual interrogation of the Google Maps timeline can be used as a first step to examine if any location data is collected and to support further examination of raw JSON data. The timeline can work as a motivating factor to demand legal access to the Google account and further raw data extraction.

4. Limitations

Manual data extraction in the application is impossible if the mobile device screen is damaged and the device is not interactive [65]. Google account can be active on several mobile devices or computers, and data modification can be done from another device while the forensic process occurs. Therefore, data must be captured as soon as the device is seized. The forensic investigator interacts with the application directly, and data modification errors can occur that should be documented during the interrogation. Actual raw data points cannot be seen, and only the approximate timing of location visits that Google believes reflects the truth are available. Data attribution is an issue because the Google account could have been used or the mobile device could have been carried by another person.

5.2 Logical data acquisition

Logical data acquisition is one level above manual acquisition. This method requires additional equipment and training. The connection between the mobile device and the forensic workstation is created using a cable or wireless connectivity, such as Wi-Fi or Bluetooth [65]. The forensic workstation contains specialized software that assists in the logical acquisition process. For this research, two fundamental techniques of logical data extraction from iOS and Android operating systems are used: Android Debug Bridge (ADB) and iPhone Backup Extractor. Appendix 3 presents the steps completed to perform ADB and iPhone Backup Extractor. A logical data extraction flow from the Google Maps timeline to the mobile device and a forensic workstation is presented in Figure 37.

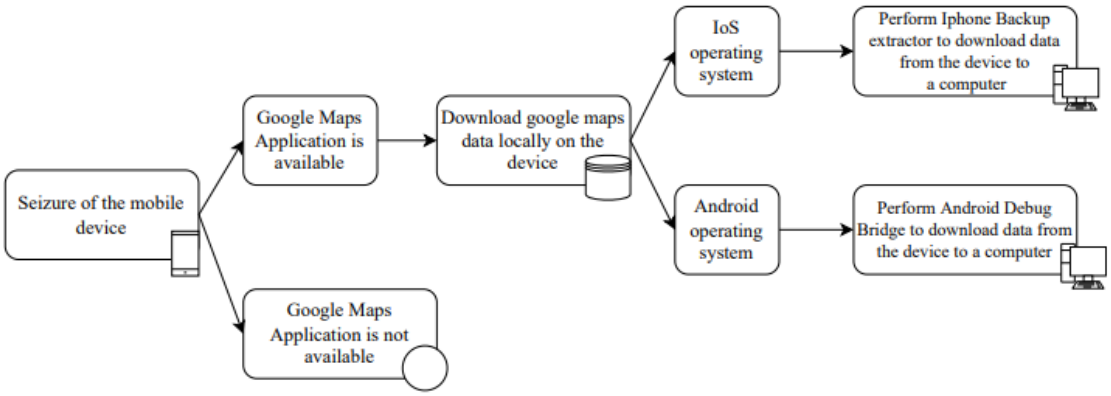


Figure 37. Logical data acquisition process flow.

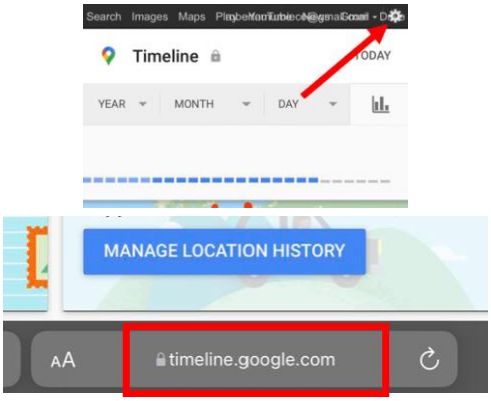
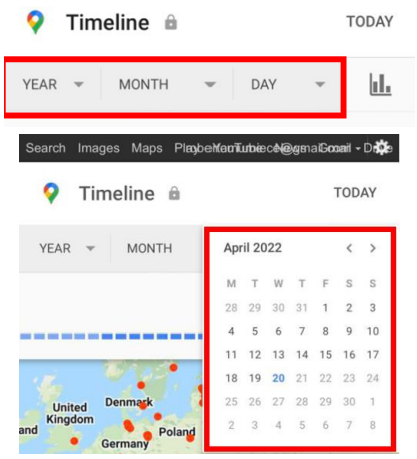
The first part of the diagram depicts data downloaded locally on the device, and the second part depicts data extraction from the downloads folder to an external forensic workstation. The forensic investigator can select an appropriate logical acquisition method depending on technical availability and knowledge. The main requirement is that during the data extraction from the mobile device to a forensic workstation, the data is maintained in its original form, supporting throughout the chain of custody and integrity.

KML geographic annotation data extraction

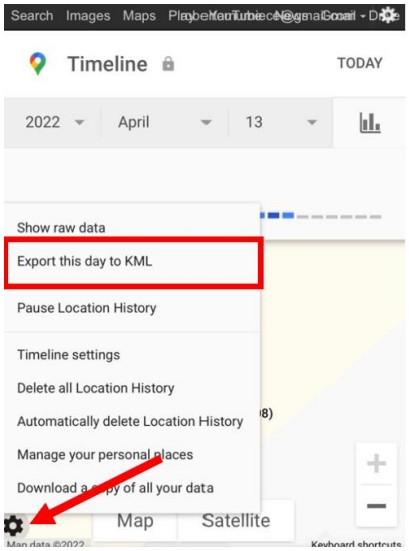
1. Data acquisition

The KML data can only be downloaded from the Google Maps timeline web page and not directly from the application. Forensic investigators must open timeline.google.com in the browser of the device of interest and export geographic annotation data from the timeline. KML data extraction is presented in Table 27.

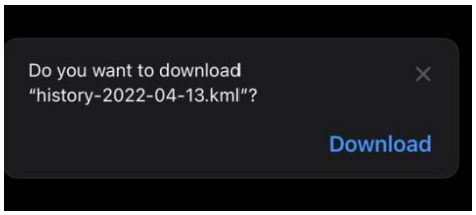
Table 27. Acquiring geographic annotation data from the timeline.

1. Open timeline.google.com in the browser of the device of interest and log in to the Google account if required 	2. Select day or period of interest 

3. Select “Export this day to KML.”



4. Download the files locally on the device downloads folder



5. Perform data acquisition from the mobile device downloads folder to a forensic workstation

2. Data interrogation

Geographic annotation of location data from the Google Maps timeline can be exported using KML files. The forensic investigator must know how to transform and understand these KML files. Acronym KML stands for Keystone Markup Language. KML files can be displayed in various geographic data browsers, like Google Earth, Adobe Photoshop, and ESRI ArcGIS Explorer [67, pp. 53–55]. This file presents the same travels and visits data as the Google Maps timeline with additional raw data points. Furthermore, this data set presents the sequence number of the event and the travel type or visited location category. A logical timeline can be recreated from this data in any other system.

3. Accuracy

This acquisition level is the first step of data acquisition, where the data is extracted from the application to a forensic workstation. The forensic value of the extracted data is above manual acquisition because raw data points, even those not in-line with the route or visit, are presented in the KML file. The raw data point absence is the major flaw in the manual acquisition level. As observed in the experiments, the accuracy of the raw data points varies based on the enabled navigation sensors. With all navigation sensors enabled, the average Google accuracy for these raw data points is 29 meters, with the average haversine distance to the ground truth even more precise at 22 meters. Raw data points hold higher forensic value than the already generated timeline. Detected visited location accuracy depends on two factors: navigation sensors and the visit type. KML export only

shows visits that were detected by the application. However, there can be no assurance that a visit will be detected by the Google Maps timeline. Only visits with probabilities above a certain threshold are presented in the Google Maps timeline. Even if the location probability is low, the visit probability must be above 60 points to be presented to the user in the application. Visit detection is also based on the unique features of the Google account, therefore the outcome can be different for two different Google accounts. Accuracy depends on the environment and enabled navigation sensors. If a visit is detected in the timeline and is in this export, then there is a high likelihood that the visit occurred, or the user was somewhere near the visited location.

4. **Use cases**

The view of the timeline provided by Google Maps is a product of the applied algorithms to the raw data points. KML data can be used to gain an insight into the underlying data of the timeline. However, this export does not present all the collected data, only particular data points that Google algorithm has linked with the timeline. This extraction layer is beneficial when the forensic investigators want to recreate the timeline by themselves and want to see the raw location data points within the timeline. Additionally, the data can be analyzed and presented in another system that may suit better for specific cases.

5. **Limitations**

The major limitation of this acquisition layer is that it requires knowledge of how to work with KML files. Additional software is needed to analyze and present the KML files, that can create additional questions about the integrity of the data modification. Google algorithm may calculate and present a wrong route or a visit. With the KML export it is not possible to assure the correctness of the visits and routes. Log-in access to the Google account may be required to download the data from the web browser of the device of interest.

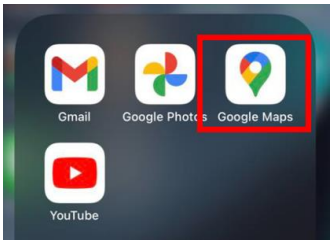
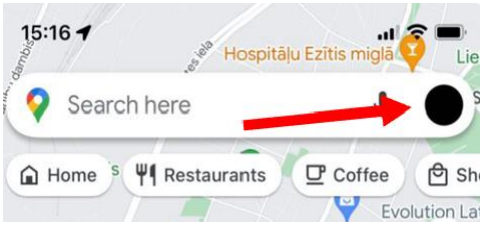
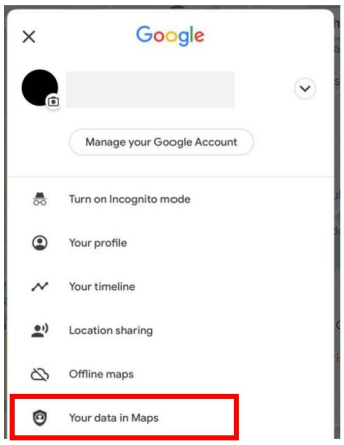
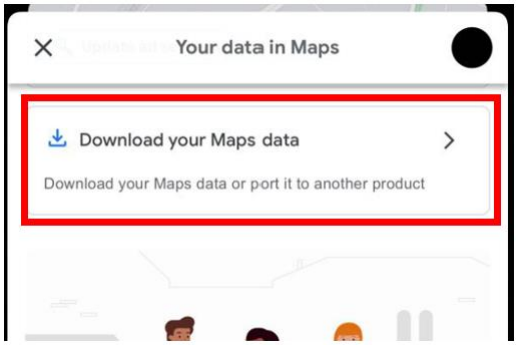
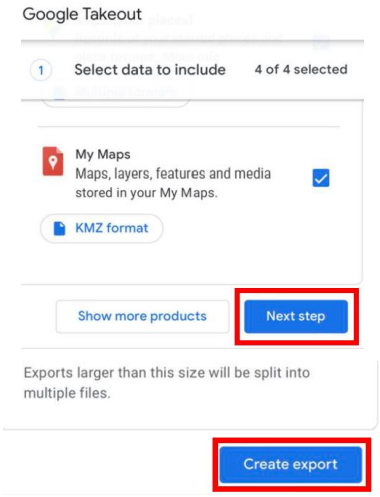
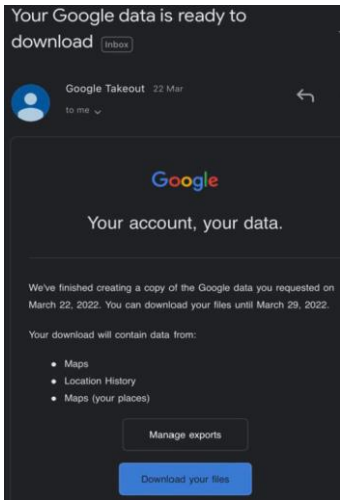
JSON raw location and semantic location history data

1. **Data acquisition**

Location history and semantic location history data are exported separately from Google Maps using the Google Takeout service. This service supports exporting raw location

history and visited place history data as presented in Table 28. The raw data contains raw physical location data points in records JSON file and semantic location history in separate JSON files for each month. The data can be extracted directly from the Google Maps application.

Table 28. Exporting raw location history from the Google Maps application.

<table>
<tr>
<td>1. Open Google Maps application on the device of interest
 
</td>
<td>2. Select the logged in Google account
 
</td>
</tr>
<tr>
<td>3. Select “Your data in Maps”
 
</td>
<td>4. Scroll down and select “Download your maps data”
 
</td>
</tr>
<tr>
<td>5. Select all maps data and “Create export”. Step may require account password.
 
</td>
<td>6. Download the files directly on the device downloads folder from the received e-mail.
 
</td>
</tr>
</table>

2. Data interrogation

This data acquisition layer requires the highest level of data transformation and manipulation. All data is in raw format, and specific skills to read and transform JSON files are required. This dataset presents all physical location data points with respective accuracies and timestamps collected by the application. The source of the raw location data point can be retrieved from the raw data export. Review the source of each location point, it will suggest the accuracy level. Cell tower source holds the highest average accuracy of 600 meters, following Wi-Fi and GPS. A timeline can be recreated by placing the latitude and longitude points on a map and joining them based on timestamp. Distinctive clusters of raw data points may indicate a visit. Comparing these clusters with semantic location history can expose visits that are not detected by the Google Maps timeline.

Semantic location history files present visited place name, address, physical location, and duration. In addition to the visited places, semantic location history files provide confidence values of each visit that must be considered when analyzing the data. Observe all visits, with the respective raw data points. Note the visit and location confidence values. Higher confidence values indicate higher likelihood of the event. Compare the visit data with raw location data, to support the insight of both datasets. A combination of raw data and visited locations is forensically valuable evidence that can be used as a primary source of proof. A complete logical timeline of movements can be recreated with both datasets, and accuracy can be reviewed for each raw data point separately, making the evidence analysis trustworthy and complete.

3. Accuracy

This extraction layer is the most forensically valuable source of location data from the Google Maps timeline. The accuracy of raw data points is available, therefore providing insight into the extent to which the forensic investigator should rely on each data point. Even if the navigation sensors are enabled or disabled at different times, the source of the location data shows the specific point's respective accuracy. As observed in the experiments, Google Maps is more pessimistic about location accuracy, and the actual accuracy is even more precise than what is suggested. Forensic investigators can expect that the device was in the radius of the Google’s suggested data point’s accuracy.

Visited places can be observed in the semantic location history files or as clusters of raw data points. Various location types are less likely to be detected, such as rural areas and only address locations; hence, other not detected visits can be discovered by reviewing the raw data point locations. Additionally, the detected semantic location accuracy can be examined using a location confidence value. The location confidence for false-positive visits is significantly lower than for visits that were detected correctly.

4. Use cases

Logical data extraction from the Google Maps timeline is performed when the forensic investigator has access to the Google account. Additionally, this extraction layer requires time and knowledge, therefore is only valid in cases when time is not a substantial constraint. This level of data acquisition can be used to support previous level findings at a later stage of the investigation. For example, if a visit is detected in the Google Maps timeline in the manual acquisition process, the finding can be supported by raw data point and semantic location history analysis.

5. Limitations

Additional software is needed to securely transfer the downloaded information from the device to a forensic workstation. The forensic investigator must know how to transform and understand JSON location files and the meaning of the various fields in these files. This technique allows a detailed investigation of raw location data points retrieved by the device and reviews all data points associated with an activity or a visit. This level of location data investigation requires time and resources. All data needs to be validated because there is high risk of improper data manipulation and presentation that can disregard the data validity. Since data is stored in remote cloud servers, issues may arise in getting access to the cloud account in time.

6 Conclusions and future research

This research has helped us to understand the forensic value of the Google Maps timeline data. We have first analyzed and understood the location data retrieved from the Google Takeout service. Additionally, we have performed ad-hoc experiments to gather a general understanding of the application's performance. Second, we have performed two types of controlled experiments to evaluate navigation sensor impact on data and place visit identification accuracy. Based on the observed phenomena we have developed a four-step acquisition protocol for location history data collected and presented in the Google Maps timeline. The key findings of the research are summarized in four points:

1. Based on the navigation sensor accuracy experiments we observed that the enabled navigation sensors directly impact the collected raw location data accuracy and the precision of the timeline. Overall, Google's accuracy is more pessimistic than the actual precision of the raw locations. The experiment results support the research [3] where four device configurations are tested (2G, 3G, Wi-Fi and GPS). In our research we have tested location services, mobile network, Wi-Fi/Wi-Fi scanning and Bluetooth/Bluetooth scanning navigation sensors.

2. The outcome of visit detection accuracy experiments shows that visit types in the city and with a semantic meaning hold a significantly higher probability of being detected by the Google Maps timeline. Rural area visits and those with only physical or address locations are not presented in the timeline after 10 minutes of dwell time. As mentioned in [20], the personal map concept can include individually significant places and routes. Therefore, visit detection accuracy can only be generalized to a certain extent because each user account has individual characteristics that may assist the application in detecting a particular location.

3. Location-based applications and Google Maps timeline are highly valuable sources of evidence from mobile devices. Majority of prior research [7], [8], [26], [31], [9] focus on location data stored locally on the mobile device and less in the cloud accounts. As mentioned in [27], there is additional complexity in analyzing

mobile cloud applications because the cloud and mobile device's local environments are separated. Our proposed data acquisition protocol only interacts with the data stored in the Google account and the cloud environment. The acquisition process consists of manual and logical evidence acquisition, and each layer holds certain data interrogation steps, use cases, accuracy, and forensic value.

4. Publicly known cases and remarks in [2], [32] as well as digital evidence admissibility standards mentioned in [34]–[39], justifies the importance of throughout Google Maps timeline examination before it can be used as a data source for an expert report in court. We have observed limited research to date about data acquisition from widely used mobile device applications and the respective cloud accounts. The research we have performed, and the proposed acquisition protocol informs the forensic community about the validity of the data, limitations that must be considered and appropriate handling procedures.

Future research

Primarily, this research should be extended with more devices as we have only analyzed three. Our validation data was collected on Samsung Galaxy S21, influenced by device-specific navigation sensors.

Secondly, each Google account holds individual features that influence the timeline. It would be necessary to further investigate the actions online or in the application, and the user attributes, such as gender or age, affect on the Google Maps timeline. Google Maps performance and collected data can be tested in an active condition, when navigation is working. Google states that the user search history can affect the visit recommendations in the Google Timeline [63]. There are no available research that confirms how often the data is modified based on the user search history. Future research could analyze how frequently this occurs in practice.

Thirdly, we have observed that there is missing academic literature on the accuracy of different location-based applications in the context of digital evidence extraction. We believe that it is integral to test and review more applications in the future, such as Apple Maps. These types of research would benefit the forensic community, as the export reports would be less challenged in the court because of extensive academic research.

References

[1] "Leading mapping apps in the United States in 2021, by downloads," *Statista*, Feb. 2021. [Online]. Available: https://www.statista.com/statistics/865413/most-popular-us-mapping-apps-ranked-by-audience/. [Accessed: Nov. 19, 2021]

[2] R. Gavin, "Judge blocks Google evidence from Troy murder trial," *Times Union*, Oct. 27, 2017. [Online]. Available: https://www.timesunion.com/news/article/Google-evidence-tossed-from-Troy-suitcase-murder-12311986.php. [Accessed: Apr. 02, 2022]

[3] A. M. Rodriguez, C. Tiberius, R. van Bree, and Z. Geradts, "Google timeline accuracy assessment and error prediction," *Forensic Sciences Research*, vol. 3, no. 3, pp. 240–255, Jul. 2018, doi: 10.1080/20961790.2018.1509187.

[4] B. Ciepłuch, R. Jacob, P. Mooney, and A. C. Winstanley, "Comparison of the accuracy of OpenStreetMap for Ireland with Google Maps and Bing Maps," *Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resuorces and Enviromental Sciences*, p. 337, Jul. 2010.

[5] F. Gülgen and B. Kiliç, "Accuracy and Similarity Aspects In Online Geocoding Services: A Comparative Evaluation For Google and Bing Maps," *International Journal of Engineering and Geosciences*, vol. 5, no. 2, pp. 109–119, Jun. 2020, doi: 10.26833/ijeg.629381.

[6] J. Sammons, *Digital Forensics: Threatscape and Best Practices*. Syngress, 2015.

[7] M. Chernyshev, S. Zeadally, Z. Baig, and A. Woodward, "Mobile Forensics: Advances, Challenges, and Research Opportunities," *IEEE Security Privacy*, vol. 15, no. 6, pp. 42–51, Nov. 2017, doi: 10.1109/MSP.2017.4251107.

[8] J. Moore, I. Baggili, and F. Breitinger, "Find Me If You Can: Mobile GPS Mapping Applications Forensic Analysis & SNAVP the Open Source, Modular, Extensible Parser," *Journal of Digital Forensics, Security and Law*, vol. 12, no. Article 7, 2017, doi: 10.15394/jdfsl.2017.1414.

[9] N.-A. Le-Khac, M. Roeloffs, and T. Kechadi, "Forensic Analysis of the TomTom Navigation Application," in *Advances in Digital Forensics X*, Berlin, Heidelberg, 2014, pp. 267–276, doi: 10.1007/978-3-662-44952-3_18.

[10] "Geocoding Service," *Google Maps Platform*. [Online]. Available: https://developers.google.com/maps/documentation/javascript/geocoding. [Accessed: Jan. 08, 2022]

[11] "Geocoding | API," *Mapbox*. [Online]. Available: https://docs.mapbox.com/api/search/geocoding/. [Accessed: Feb. 26, 2022]

[12] J. Hightower and G. Borriello, "Location systems for ubiquitous computing," *Computer*, vol. 34, no. 8, pp. 57–66, Aug. 2001, doi: 10.1109/2.940014.

[13] O. Kounadi, T. J. Lampoltshammer, M. Leitner, and T. Heistracher, "Accuracy and privacy aspects in free online reverse geocoding services," *Cartography and Geographic Information Science*, vol. 40, no. 2, pp. 140–153, Mar. 2013, doi: 10.1080/15230406.2013.777138.

[14] C. A. Davis Jr. and R. O. de Alencar, "Evaluation of the quality of an online geocoding resource in the context of a large Brazilian city," *Transactions in GIS*, vol. 15, no. 6, pp. 851–868, 2011, doi: 10.1111/j.1467-9671.2011.01288.x.

[15] H. J. Miller, “Tobler’s First Law and Spatial Analysis,” *Annals of the Association of American Geographers*, vol. 94, no. 2, pp. 284–289, Jun. 2004, doi: 10.1111/j.1467-8306.2004.09402005.x.
[16] Q. Li, Y. Zheng, X. Xie, Y. Chen, W. Liu, and W.-Y. Ma, “Mining user similarity based on location history,” in *Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems - GIS ’08*, Irvine, California, 2008, pp. 1–10, doi: 10.1145/1463434.1463477.
[17] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, “Mining interesting locations and travel sequences from GPS trajectories,” in *Proceedings of the 18th international conference on World wide web - WWW ’09*, Madrid, Spain, 2009, pp. 791–800, doi: 10.1145/1526709.1526816 [Online]. Available: http://portal.acm.org/citation.cfm?doid=1526709.1526816. [Accessed: Jan. 08, 2022]
[18] D. Ashbrook and T. Starner, “Using GPS to learn significant locations and predict movement across multiple users,” *Personal and Ubiquitous Computing*, vol. 7, no. 5, pp. 275–286, Oct. 2003, doi: 10.1007/s00779-003-0240-0.
[19] X. Cao, G. Cong, and C. S. Jensen, “Mining significant semantic locations from GPS data,” *Proc. VLDB Endow.*, vol. 3, no. 1–2, pp. 1009–1020, Sep. 2010, doi: 10.14778/1920841.1920968.
[20] L. Liao, D. J. Patterson, D. Fox, and H. Kautz, “Building Personal Maps from GPS Data,” *Annals of the New York Academy of Sciences*, vol. 1093, no. 1, pp. 249–265, 2006, doi: 10.1196/annals.1382.017.
[21] C. Zhou, N. Bhatnagar, S. Shekhar, and L. Terveen, “Mining Personally Important Places from GPS Tracks,” in *2007 IEEE 23rd International Conference on Data Engineering Workshop*, Apr. 2007, pp. 517–526, doi: 10.1109/ICDEW.2007.4401037.
[22] B. Martini, Q. Do, and K.-K. R. Choo, “Mobile Cloud Forensics: An Analysis of Seven Popular Android Apps,” *The Cloud Security Ecosystem*, Jun. 2015, doi: 10.1016/B978-0-12-801595-7.00015-X.
[23] N. Al Mutawa, I. Baggili, and A. Marrington, “Forensic analysis of social networking applications on mobile devices,” *Digital Investigation*, vol. 9, pp. S24–S33, Aug. 2012, doi: 10.1016/j.diin.2012.05.007.
[24] P. Sharma, D. Arora, and T. Sakthivel, “Enhanced Forensic Process for Improving Mobile Cloud Traceability in Cloud-Based Mobile Applications,” *Procedia Computer Science*, vol. 167, pp. 907–917, Jan. 2020, doi: 10.1016/j.procs.2020.03.390.
[25] H. Zhang, L. Chen, and Q. Liu, “Digital Forensic Analysis of Instant Messaging Applications on Android Smartphones,” in *2018 International Conference on Computing, Networking and Communications (ICNC)*, Mar. 2018, pp. 647–651, doi: 10.1109/ICCNC.2018.8390330.
[26] A. Levinson, B. Stackpole, and D. Johnson, “Third Party Application Forensics on Apple Mobile Devices,” in *2011 44th Hawaii International Conference on System Sciences*, Jan. 2011, pp. 1–9, doi: 10.1109/HICSS.2011.440.
[27] N. Samet, A. Ben Letaïfa, M. Hamdi, and S. Tabbane, “Forensic investigation in Mobile Cloud environment,” in *The 2014 International Symposium on Networks, Computers and Communications*, Jun. 2014, pp. 1–5, doi: 10.1109/SNCC.2014.6866510.
[28] S. Zargari and D. Benford, “Cloud Forensics: Concepts, Issues, and Challenges,” in *2012 Third International Conference on Emerging Intelligent Data and Web Technologies*, Sep. 2012, pp. 236–243, doi: 10.1109/EIDWT.2012.44.

[29] S. Almulla, Y. Iraqi, and A. Jones, “A State-Of-The-Art Review of Cloud Forensics,” *JDFSL*, vol. 9, no. 4, Article 2, 2014, doi: 10.15394/jdfsl.2014.1190.
[30] A. Edens, *Cell Phone Investigations: Search Warrants, Cell Sites and Evidence Recovery*. Police Publishing, 2014.
[31] G. M. Jones and S. G. Winster, “Forensics Analysis On Smart Phones Using Mobile Forensics Tools,” *International Journal of Computational Intelligence Research*, vol. 13, no. ISSN 0973-1873, p. 12, 2017.
[32] J. Valentino-DeVries, “Tracking Phones, Google Is a Dragnet for the Police,” *The New York Times*, Apr. 13, 2019 [Online]. Available: https://www.nytimes.com/interactive/2019/04/13/us/google-location-tracking-police.html. [Accessed: Apr. 02, 2022]
[33] E. Casey and C. W. Rose, *Handbook of Digital Forensics and Investigation*. San Diego: Academic Press, 2010 [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780123742674000021. [Accessed: Oct. 31, 2021]
[34] R. McKemmish, “When is Digital Evidence Forensically Sound?,” in *Advances in Digital Forensics IV*, Boston, MA, 2008, pp. 3–15, doi: 10.1007/978-0-387-84927-0_1.
[35] E. Casey, *Digital Evidence and Computer Crime: Forensic Science, Computers, and the Internet*. Academic Press, 2011.
[36] S. Vömel and F. C. Freiling, “Correctness, atomicity, and integrity: Defining criteria for forensically-sound memory acquisition,” *Digital Investigation*, vol. 9, no. 2, pp. 125–137, Nov. 2012, doi: 10.1016/j.diin.2012.04.005.
[37] S. Goodison, “Digital Evidence and the U.S. Criminal Justice System,” Apr. 2015.
[38] T. Wu, F. Breitinger, and S. O’Shaughnessy, “Digital forensic tools: Recent advances and enhancing the status quo,” *Forensic Science International: Digital Investigation*, vol. 34, no. Article 300999, Sep. 2020, doi: 10.1016/j.fsidi.2020.300999.
[39] R. Adams, V. Hobbs, and G. Mann, “The Advanced Data Acquisition Model (Adam): A Process Model for Digital Forensic Practice,” Doctoral Thesis, Murdoch University, 2012.
[40] A. Hoog, *Android Forensics: Investigation, Analysis and Mobile Security for Google Android*. Elsevier, 2011.
[41] D. DeMatteo, S. Fishel, and A. Tansey, “Expert Evidence: The (Unfulfilled) Promise of Daubert,” *Psychological Science in the Public Interest*, vol. 20, no. 3, pp. 129–134, Dec. 2019, doi: 10.1177/1529100619894336.
[42] A. Alarifi *et al.*, “Ultra Wideband Indoor Positioning Technologies: Analysis and Recent Advances,” *Sensors*, vol. 16, no. 5, p. 707, May 2016, doi: 10.3390/s16050707.
[43] R. Mautz, “Indoor positioning technologies,” Habilitation Thesis, ETH Zurich Department of Civil, Environmental and Geomatic Engineering, Institute of Geodesy and Photogrammetry, 2012.
[44] H. Liu, H. Darabi, P. Banerjee, and J. Liu, “Survey of Wireless Indoor Positioning Techniques and Systems,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1067–1080, Nov. 2007, doi: 10.1109/TSMCC.2007.905750.
[45] “Other Global Navigation Satellite Systems (GNSS),” *GPS.gov*. [Online]. Available: https://www.gps.gov/systems/gnss/. [Accessed: Apr. 21, 2022]
[46] E. D. Kaplan and C. Hegarty, *Understanding GPS/GNSS: Principles and Applications, Third Edition*. Artech House, 2017.

[47] “Global Navigation Satellite Systems (GNSS),” *United Nations Office for Outer Space Affairs*. [Online]. Available: https://www.unoosa.org/oosa/en/ourwork/psa/gnss/gnss.html. [Accessed: Apr. 21, 2022]
[48] C. Gentile, N. Alsindi, R. Raulefs, and C. Teolis, *Geolocation Techniques: Principles and Applications*. Springer Science & Business Media, 2012.
[49] “Global Positioning System,” *Oxford Advanced Learner’s Dictionary*. 2015 [Online]. Available: https://www.oxfordlearnersdictionaries.com/definition/english/global-positioning-system. [Accessed: Nov. 06, 2021]
[50] “GPS Accuracy,” *GPS.gov*. [Online]. Available: https://www.gps.gov/systems/gps/performance/accuracy/. [Accessed: Dec. 12, 2021]
[51] C. Feng, W. S. A. Au, S. Valaee, and Z. Tan, “Compressive Sensing Based Positioning Using RSS of WLAN Access Points,” in *2010 Proceedings IEEE INFOCOM*, Mar. 2010, pp. 1–9, doi: 10.1109/INFCOM.2010.5461981.
[52] *Mobile Network Guide Improving Mobile Signal*. Australia: Powertec Telecommunications Pty Ltd, 2014 [Online]. Available: https://www.mobilenetworkguide.com.au/pdf/Mobile-Network-Guide-Improving-Mobile-Signal.pdf. [Accessed: Dec. 12, 2021]
[53] L. Ezema, C. Ani, and G. Nw. Ezeh, “Mobile Location Estimation in GSM/UMTS,” *IJETR*, vol. 1, no. 3, pp. 63–70, Jul. 2019.
[54] L. A. Martínez Hernández, S. Pérez Arteaga, G. Sánchez Pérez, A. L. Sandoval Orozco, and L. J. García Villalba, “Outdoor Location of Mobile Devices Using Trilateration Algorithms for Emergency Services,” *IEEE Access*, vol. 7, pp. 52052–52059, 2019, doi: 10.1109/ACCESS.2019.2911058.
[55] F. Chen, C. Yang, W. Yu, X. Le, and J. Yang, “Research on mobile GIS based on LBS,” in *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS ’05.*, Jul. 2005, vol. 2, p. 4 pp.-, doi: 10.1109/IGARSS.2005.1525256.
[56] D. Arctur and M. Zeiler, *Designing Geodatabases: Case Studies in GIS Data Modeling*. ESRI, Inc., 2004.
[57] K.-T. Chang, “Geographic Information System,” in *International Encyclopedia of Geography*, 2019th ed., John Wiley & Sons, Ltd [Online]. Available: https://doi.org/10.1002/9781118786352.wbieg0152.pub2. [Accessed: Dec. 11, 2021]
[58] P. Bellavista, A. Küpper, and S. Helal, “Location-Based Services: Back to the Future,” *IEEE Pervasive Computing*, vol. 7, no. 2, pp. 85–89, Apr. 2008, doi: 10.1109/MPRV.2008.34.
[59] H. Huang and S. Gao, “Location-Based Services,” 2018.
[60] H. Huang, G. Gartner, J. M. Krisp, M. Raubal, and N. Van de Weghe, “Location based services: ongoing evolution and research agenda,” *Journal of Location Based Services*, vol. 12, no. 2, pp. 63–93, Apr. 2018, doi: 10.1080/17489725.2018.1508763.
[61] H. R. Schmidtke, “Location-aware systems or location-based services: a survey with applications to CoViD-19 contact tracking,” *J Reliable Intell Environ*, vol. 6, no. 4, pp. 191–214, Dec. 2020, doi: 10.1007/s40860-020-00111-4.
[62] P. Gilski and J. Stefański, “Survey of Radio Navigation Systems,” *International Journal of Electronics and Telecommunications*, vol. 61, Mar. 2015, doi: 10.1515/eletel-2015-0006.

[63] “Google Maps Timeline,” *Google Maps Help*. [Online]. Available: https://support.google.com/maps/answer/6258979. [Accessed: Jan. 17, 2022]
[64] “How Google uses location information,” *Google Privacy and Terms*. [Online]. Available: https://policies.google.com/technologies/location-data. [Accessed: Oct. 13, 2021]
[65] R. Ayers, S. Brothers, and W. Jansen, “Guidelines on mobile device forensics,” National Institute of Standards and Technology, NIST SP 800-101r1, May 2014 [Online]. Available: http://dx.doi.org/10.6028/NIST.SP.800-101r1
[66] S. C. Sathe and N. M. Dongre, “Data acquisition techniques in mobile forensics,” presented at the 2018 2nd International Conference on Inventive Systems and Control (ICISC), Jan. 2018, pp. 280–286, doi: 10.1109/ICISC.2018.8399079.
[67] M. Harrington and M. Cross, *Google Earth Forensics: Using Google Earth Geo-Location in Digital Forensic Investigations*. Syngress, 2014.

Appendix 1 Visited location detailed data matrix

Visit detection success rate %, Balvi

Nr.	Type	iPhone7 Plus	Samsung Galaxy A3	HuaweiP20 Lite	%
1	Rural	0	0	0	0%
2	Rural	0	0	0	0%
3	Rural	0	0	0	0%
4	Town	0	0	1	33%
5	Town	1	1	0	67%
6	Town	1	1	0	67%
7	Town	0	1	0	33%
8	Rural	0	0	0	0%
9	Rural	0	0	0	0%
10	Rural	0	0	0	0%
		20%	30%	10%	20%

Visit detection success rate %, Riga

Nr.	Type	iPhone7 Plus	Samsung Galaxy A3	HuaweiP10 Lite	%
1	Outside	0	0.5	0	17%
2	Outside	0.5	0	0	17%
3	Outside	0	0	0	0%
4	Inside	0.5	1	1	83%
5	Inside	0	1	1	67%
6	Inside	0.5	1	1	83%
7	Outside	0	0	0	0%
8	Outside	0.5	0	0	17%
9	Outside	0	1	1	67%
10	Inside	0	1	1	67%
		20%	55%	50%	42%

Appendix 2 Google Maps permissions on a mobile device

Permission name	**Collected data**	**Available for Google Maps application**
Body sensors	Sensor information about the user's body	No
Calendar	Usage of the default calendar	No
Call logs	Call history	No
Camera	Use a camera to take pictures	Yes
Contacts	See contact list	Yes
Location	Request device's location	Yes
Microphone	Record audio	Yes
Nearby Bluetooth devices	Discover and connect to nearby devices	Yes
Phone	Make and manage phone calls	No
Physical activity	Request information and recognize physical activity types	Yes
SMS	See and send text messages	No
Storage	Request files from the phone in storage	Yes
Files and media	Request files from the phone	Yes

Appendix 3 iPhone Backup extractor and ADB process

The iPhone Backup extractor required four steps to be performed:

1. Install and run https://www.iphonebackupextractor.com/
2. Using a cable to connect the device with the computer
3. On the device, select: “Allow this device to access photos and videos.”
4. Open the downloads folder and download the files locally on the computer

Android Debug Bridge required five steps to be performed:

1. Download and unzip Android SDK Platform files
2. Run command prompt in the respective directory
3. Allow developer tools on the mobile device and select USB debugging
4. Connect the device with the computer with a USB cable and switch to the file transfer mode
5. Save files from the downloads folder from the phone to the computer with the *“./adb pull”* command