PRIVACY-PRESERVING VEHICLE MILES TRAVELED (PPVMT) TAX

BY

GAURAV LAHOTI

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Computer Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Adviser:

Professor Carl A. Gunter

Abstract

Motor fuel taxes form a great part of the total revenue collected for the development and maintenance of surface transportation. Gasoline tax payments are becoming a matter of concern as the share of Battery Vehicles (BVs) increases in the market. Those funds need to be collected in some other way in light of the decreased gasoline sales.

In the recent years, a concept of mileage-based tax, called Vehicle Miles Traveled (VMT) tax, has been developed to address this concern. This approach calculates tax by monitoring vehicle road usage through GPS and odometer data. GPS data determines vehicle's region and the rate per mile to be assessed. The number of miles driven is taken from the odometer.

Collection of fine-grained GPS data is privacy invasive and has been a strong reason against adoption of VMT tax. Coarsening the location dependent data does not help either. Ensuring secure computation of the data, and validation of GPS signals also of prime concern as the user might tamper with the system to report less miles.

We propose Privacy-Preserving Vehicles Miles Traveled (PPVMT) tax based on additive secret-sharing to solve the privacy issues. In our model, the car computes the total miles driven in each (tax jurisdiction) region. The car splits the total miles of each region into random looking numbers which can later be aggregated in a specific manner to determine tax owed by each user and tax share of each region.

We also propose 'Car-as-a-Smartphone' model which reasons that features available in a car in the near future will be similar to a current smartphone. To detect system tampering and signal spoofing, we propose validation of the untrusted data from GPS and odometer with inertial motion sensors. We implemented a technique to verify the pattern of location coordinates reported by the GPS in the car from the gyroscope data. The technique raises the cost and the skill required to tamper with the system.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The price for motor fuel at gas stations consists of the commodity price plus taxes charged by federal, state and, in certain cases, local governments. These taxes are used for the maintenance and improvement/expansion of the surface transportation infrastructure for vehicular traffic. The deepening penetration of Battery Vehicles (BVs), which use the 'fuel' from the power grid, threatens to reduce such funding substantially.

Covering up the revenue deficit can be done in many different ways.

- One strategy would be to charge BV buyers a high registration fee at the time of vehicle purchase. This will be an impractical and unfair mechanism as some cars are driven significantly more than others. A person driving less will end up subsidizing the cost for the person who drives more. Besides, some cars function for more years than others. Such cars use the infrastructure more than other cars leading to another inequality.

- Taxing car batteries to recover the cost can be another strategy but one can argue about its practicality and fairness. Each battery is functional for different length of time. Generally, batteries last for long time and to cover the cost of infrastructure usage at the time of purchase would require charging customers a high amount of tax. It will be unfair as some batteries might not last long.

- Taxing the power used for charging a BV from electric grid can also be a potential solution but it is hard to implement effectively. Currently, many BVs can be charged from domestic sockets and implementing this system will require standardization and regularization of electricity transfer from grid to car. Also, laws would need to be setup to transfer the money collected to the transportation authorities. But

the major issue is that the system will be easy to circumvent by custom equipment or unauthorized personnel. Converting electricity to charge the batteries requires modest skills that many people possess. Consumers can utilize the electricity meant for household purposes for charging their cars which will make be difficult to monitor and collect the tax for the electricity consumed for charging cars.

- Taxing vehicles proportionately to the number of miles driven, commonly known as Vehicle Miles Traveled (VMT) tax, seems to be the most reasonable choice as it taxes motorists based on their use of surface transportation infrastructure. This technique is fair to the motorists as they pay on the basis of their usage. On top of this, the VMT tax is not dependent on the type of fuel used. If we get hydrogen or any other fuel car in the future, this tax can still collect the correct tax amount.

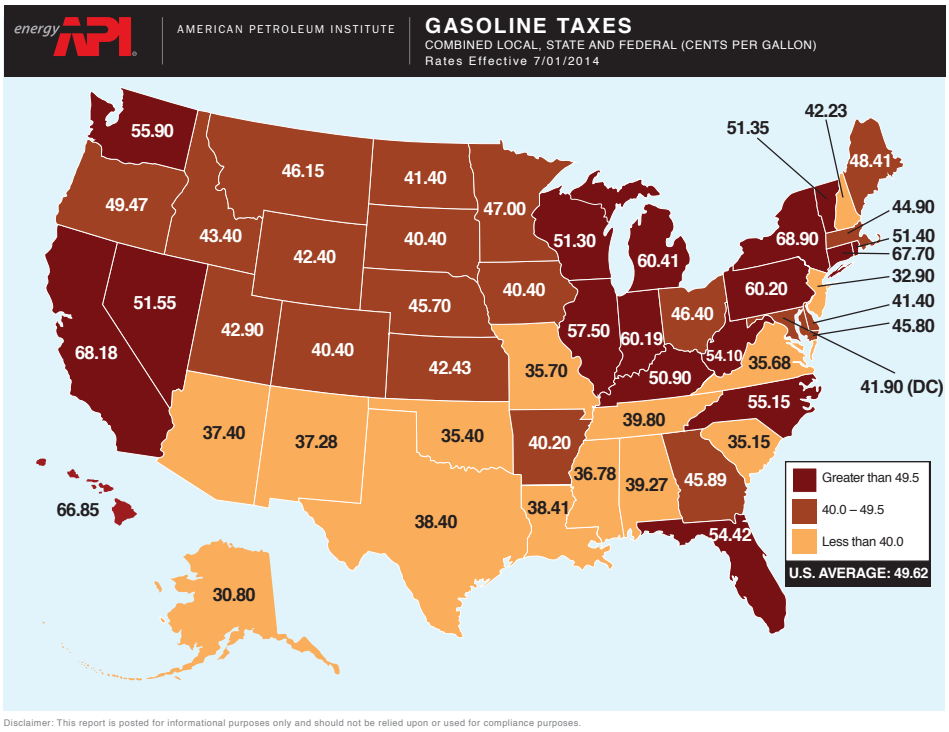


Figure 1.1: Gasoline Tax Map (Source: American Petroleum Institute)

VMT tax has two variants: one monitors location of the vehicle (mostly using GPS) and the other does not. Monitoring location allows different

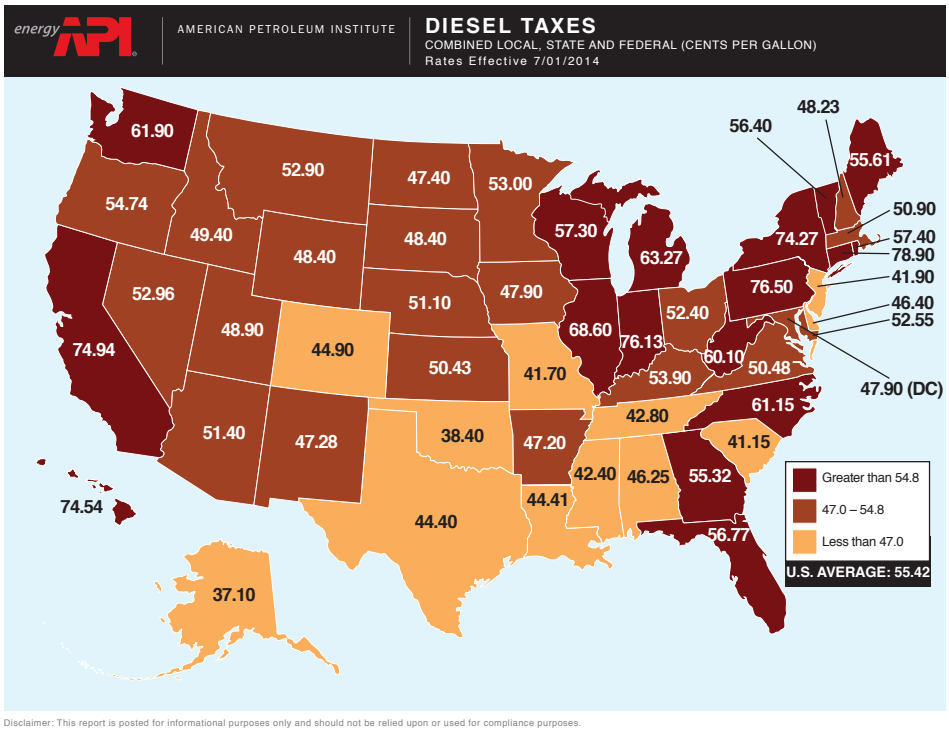


Figure 1.2: Diesel Tax Map (Source: American Petroleum Institute)

geographical regions to have different tax rates. Every state (and many regional authorities [1]) charges their own tax on the motor fuel on top of the tax collected by federal government. The federal government charges tax at 18.3 cents per gallon for gasoline and 24.3 cents per gallon for diesel [2]. Figure 1.1 maps total gasoline taxes and Figure 1.2 maps total diesel taxes of 2014 in each state. The gasoline prices vary by more than 100% between Alaska and New York.

Location recording, however, is privacy invasive and is a major obstacle to the acceptance of this tax. It has been shown that location traces can be used to infer personal information and activities ([3, 4]). For instance, location traces of a person can be used to infer the address of the house and the office. Knowing this, a web search on the social media or websites, which allow searching for people based on their personal information, can reveal more information about the person. Similarly, the data can also be used to find the visits to a hospital, vacation spots, grocery store, and many other personal preferences. Finding the number of visits to hospitals can be

privacy invasive as it may reveal the disease. For instance, if the visits are very frequent, and if it is a cancer hospital, it is safe to conclude that the person or some other family member has cancer.

To protect the privacy of the user, we propose *Privacy-Preserving Vehicles Miles Traveled (PPVMT)* tax. In this, we provide the benefit of location tracking in the VMT tax, viz. the ability to determine the tax amount of a region, without invading the location privacy of the user. We accomplish it by computing the total miles driven in each (tax jurisdiction) region locally on the car. When it is time to upload the information for tax computation, the car splits the total miles accumulated in each region into random looking numbers whose sum is the mileage in that region. The splits are sent to non-colluding servers which perform aggregation on data coming from a potentially large number of cars in a specific manner. The servers, even if they have all the information about the car, cannot infer any information about the location or the miles driven in each region. The aggregated result from all the servers is sent to the authority responsible for taxing users and allocating the tax amount to each government. The data is again aggregated by the authority in a specific manner to determine tax owed by each user and tax share of each region.

Not all of the cars today support the processing power required to perform the local computation. Some cars come with processing power to run custom applications, in-built Internet connectivity, GPS and navigation. Such cars should be able to do the calculation and send the data to the servers. For cars incapable for doing the calculation, installing an external device, either a on-board computer or a OBD-II (Onboard Diagnostic) compatible device, can work.

Cars manufacturers are constantly coming with new features and, with that in mind, we propose 'Car-as-a-Smartphone' model which reasons that features available in a car in the near future will be similar to a current smartphone. These cars will be able to run third party applications and provide powerful processing power. In such an environment, the tax application can run as an application offered by the government.

A user has the incentive to tamper with the system to lower his tax amount. The current generation of cars is highly insecure and vulnerable to exploits. [5] has shown that it is possible to compromise even the critical functions in a car, such as the brakes of the car, and overwrite the software

in the car with a custom software without raising any suspicion. In VMT tax scenario, the tampering can be done at different stages:

1. GPS signals may be blocked by jamming or the receiver may be enclosed in a Faraday cage, resulting in no data.

2. The odometer may be compromised to not report any distance covered.

3. The software performing the computations may be compromised such that it lowers the tax amount

4. Data coming from GPS receiver or odometer may be modified by some custom middle hardware installed by the user.

5. GPS signals can be spoofed such that the receiver reports false locations to lower the tax amount. The falsified location coordinates may lie in a tax jurisdiction with lower tax rate per mile.

The first two are easy to detect with inertial motion sensors. If the GPS or the odometer does not respond but the motion sensors get measurements of driving, it may be a case of some fault or tampering. PPVMT assumes that car manufacturers take steps to prevent or detect the third scenario. PPVMT provides tamper resistance to tackle the fourth and fifth scenario.

GPS and odometer readings can be spoofed as value recorded depends on external electromagnetic signals. Though there are techniques to detect spoofing, there is no guarantee that a more sophisticated attack can be detected or prevented. We believe that using inertial motions sensors, which provide measurements from the mechanics of a moving car, can validate the GPS location trace and odometer data to detect tampering. Accelerometer and gyroscope data can be compared with timestamped GPS and odometer data to match the characteristics of the car movement.

We implemented a technique to verify the pattern of location coordinates reported by the GPS in the Car from the gyroscope data. Gyroscope, which measures angular velocity, reports measurements when the car takes a turn. This turn can be compared with the GPS data to determine if something is amiss. An Android app has been developed as a proof-of-concept with good results.

1.1 Objective

This thesis was developed with the following goals:

- Recognize the problem of decreasing revenue for surface transportation infrastructure development and maintenance because of reduced gasoline sales
- Develop requirements for any Vehicle Miles Traveled (VMT) tax solution which can allocate the tax to the governments based on the miles driven by vehicles in that region
- Design a solution for VMT tax, satisfying the requirements, such that is protects the location privacy of the motorists
- Incorporate tamper resistance in the design to prevent users to from decreasing their tax amounts
- Develop a proof-of-concept of the proposed design

1.2 Contribution

Contributions of the thesis include:

- Developed functional, performance, security and policy requirements for a VMT tax solution, acknowledging the entities that will get affected by a VMT tax
- Proposed an architecture to calculate VMT tax and allocate the tax amount proportionately to all the jurisdictions
- Incorporated location privacy-protection in the design to prevent any leakage of personal information
- Proposed a tamper-resistance model which validates untrusted GPS and odometer data by checking the values against inertial motion sensor measurements
- Implemented a proof-of-concept of privacy-preserving technique and tamper resistance on Android by validating GPS data against gyroscope measurements

- Conducted field tests to collect GPS and gyroscope data and to develop a model for verifying genuineness of GPS data using gyroscope data

1.3 Organization

The rest of the thesis is organized as follows. Chapter 2 describes the motivation behind the problem and the background required to follow rest of the thesis. The main motivation of decreasing revenues due to reduced gasoline sales has been discussed. It also mentions various other automobile taxation systems in place. It provides the motivation behind securing the electronic system of the current generation of cars. It explains Global Positioning System (GPS) and inertial motions sensors in Android. Chapter 3 describes the threat to motorists' privacy because of location monitoring and inference attacks. It also overviews cyber-security issues in the automobiles. Chapter 4 provides the functional, performance, security and policy requirements for a VMT tax solution. Chapter 5 proposes PPVMT tax design and the mechanism to compute taxes of each individual and the governments. It also suggests ways to provide tamper resistance. Chapter 6 proposes 'Car-as-a-Smartphone' model and provides details of the implementation, including the model to validate GPS data from gyroscope measurements in a car. Chapter 7 discusses on other VMT projects, Pay-As-You-Drive (PAYD) insurance, smart meters and GPS spoofing countermeasures. Chapter 8, the last chapter, concludes the thesis with a summary and potential of this technology in problems. It also outlines directions for future work.

Chapter 2

Motivation and Background

This chapter presents some financial figures which are the main motivation for the governments to pursue VMT tax. It also presents relevant information about GPS, inertial motion sensors and the current state of automobile cybersecurity. The background information on these topics will help reason about the design and implementation decisions which appear later in the thesis.

2.1 Automobile Taxation

The revenue for surface transportation infrastructure has reached really low as compared to outlays. One example of deficit is Highway Trust Fund (HTF) administered by United States Department of Transportation. This fund receives the federal part of the tax from fuel sales. It is the primary financing mechanism for the nation's surface transportation system. According to Congressional Budget Office (CBO), cash outflows have outpaced tax revenues in HTF since 2001 and revenue from gasoline tax has declined over the 2010-2013 period. In the 2008-2014 period, $54 billion were transferred from the general fund of the Treasury to the HTF to make up for the shortfall [6]. In its April 2014 baseline for programs funded from the highway account, CBO projects outlays of $465 billion and revenues of $343 billion from 2015 through 2024, resulting in a cumulative shortfall of about $120 billion in 2024 in its highway account (Figure 2.1). The Figure plots the receipts, outlays, and balance or shortfall of the highway account for the Highway Trust Fund under Congressional Budget Office's April 2014 baseline.

VMT tax has been deployed worldwide but is commonly applied to only trucks. New Zealand applies it to all heavy vehicles and cars running on untaxed fuel (includes diesel and excludes petrol, CNG, and LPG) [7]. The

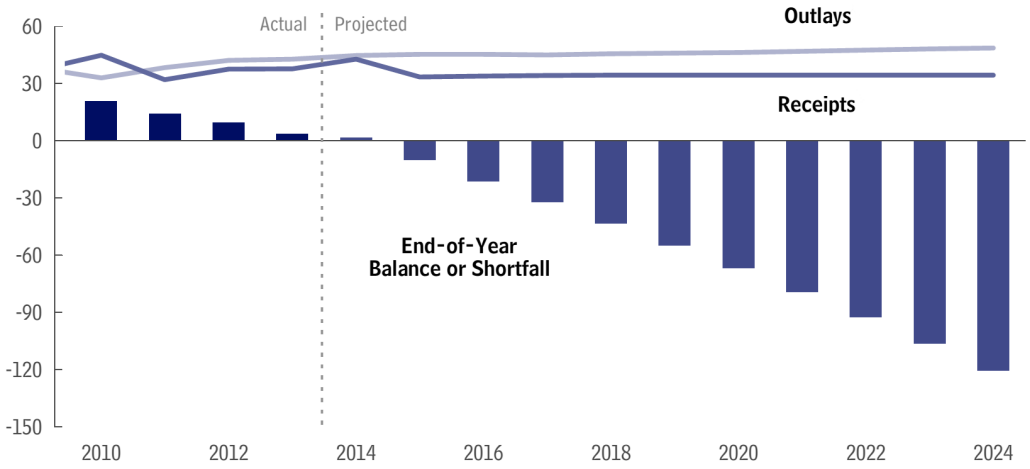


Figure 2.1: Highway Account of Highway Trust Fund (in Billions of Dollars)(Source: http://www.cbo.gov/publication/45416)

tax has been proposed in many states in USA and but has not been enforced on general public. Oregon has taken the lead by passing Senate Bill 810 [8]. It is a voluntary program for up to 5,000 motorists and becomes operational from July 1, 2015. Participants will be charged a rate of 1.5 cents per mile and 50% of the money collected will be allocated to the Department of Transportation, 30% to counties and 20% to cities irrespective of miles driven in any county or city. The bill has been made for VMT tax without location monitoring and does not enforce a location monitoring mechanism.

Currently, there is no standardized VMT tax collection mechanism for personal vehicles. One technique can be to appoint a set of authorized entities which manually inspect the car and record odometer readings. Such a mechanism will be very expensive to execute on a large scale because of the labor costs involved. Another option is to ask motorists to report the mileage from their cars to the authority or buy a pre-paid license. For example, New Zealand Transport Agency expect motorists to buy the license in 1000 km units[7] before driving the vehicle. The process is not very user friendly, wastes many man hours and trusts the people to buy a license. A popular choice has been a use of an external on-board computers (or On-Board Unit (OBU)) in the cars. Cars are modified to integrate these small computers which record the value from odometer and may have location monitoring and Internet connectivity. This mechanism requires one time installation of the unit and manual intervention only when there is an issue such as the unit

going bad. The major downside is the cost of the unit. In US alone, there are more than 250 million cars on the road and installing a $30 unit will cost $7.5 billion excluding labor and other charges. Setting up the IT infrastructure will be another cost but it should be a small fraction of the cost of the units. Cars with in-built Internet connectivity and capability of running custom software can send the relevant data to the servers automatically. Cars supporting features such as CarPlay from Apple [9], which connects the smartphone to the car and allows for data exchange, can use the smartphone to send the data to the servers. The downside of the approach is that not all cars on the road today support such system. Currently, these are available only in luxury cars and few mid-range cars.

Without location monitoring, people would end up paying taxes for driving in private areas and regions not maintained by the charging authority. For instance, volunteers paying VMT tax to a state will end up paying taxes for driving in other states. Even in a state, taxes will need to be averaged for all regions and people living in areas with lesser tax rates and lower cost of living end up subsidizing for people living in areas with higher cost of living.

2.2 Automobile Cybersecurity

Automobiles today contain Electronic Control Units (ECUs) which ensure a wide range of functionality, including engine control, brakes, lights, communication and the entertainment system. A modern automobile may contain 70-100 ECUs, running millions of lines of code [10]. Until some years ago, an automobile was a network of ECUs with no connectivity to outside world. It has changed with the introduction of wireless communication in cars. Unfortunately, the software stack running the components still uses the old, vulnerable software or its variants. Current cars have old Unix-like environment supporting electronic system of the car [5]. Since the software present on most of the cars cannot be updated without taking the car to the mechanic, vulnerabilities in the software stay with the car. With additional features coming to the cars, the number of ECUs and the amount of code, both are increasing. Cars now support GPS and navigation along with Internet connectivity.

Weak security of the car's architecture and software allows an attacker to

inject packets wirelessly to compromise car components. The weak internal security mechanism allows a compromised component to exploit other components. Bluetooth, telematics unit, and Tire Pressure Monitoring System (TPMS) have been shown to be vulnerable to such attacks [11, 12].

All cars in US since 2008 run CAN (Controller Area Network) protocol to deliver message from one component to another. It is a link layer protocol without any inbuilt security mechanism (ISO 11898 [13]). The broadcast nature of the CAN protocol allows any malicious component on the network to snoop on all the communications and send packets to other components. Absence of source identifiers or authenticator fields in the CAN protocol and a weak access control mechanism allow a non-critical compromised component to send malicious packets to critical components without getting identified. Researchers have been able to exploit and control critical and non-critical components of a car through wired OBD-II port (available in all cars in US after 1996) while the car was running [5]. Attacks through OBD-II port are difficult to execute as the port is present inside the car and requires access to the car.

There have been past projects to improve the security of the vehicle. OVERSEE (Open VEhiculaR SEcurE platform) was one such project to provide a secure, standardized and generic communication and application platform for vehicles [14]. It aimed to create a single point of access to internal and external communication channels. There is not much information about the final outcome of the project which ended in 2011.

2.3 Global Positioning System (GPS)

GPS is a satellite based navigation system which provides location information. A GPS receiver records the signal from the satellite and triangulates the location based on the information in the signals. The receiver can determine its latitude, longitude and altitude.

There are currently more than 30 GPS satellites in operation. Each GPS satellite continuously sends its location and the time at which the signal was sent. The receiver records the time at which the signal was received. Using the difference of the two times and the speed at which the signal travels, it calculates the distance of the satellite from the receiver. Using the location

and the distance from satellites, the receiver computes its position on earth.

Triangulation requires 4 or more GPS satellites in direct line of sight as the GPS signal power is weak. It makes GPS unusable in building indoors and covered parking lots. GPS triangulation requires distance of the satellite from the receiver. Location calculation is erroneous if the calculated distance has errors. This is often the case when the receiver is present among high rise buildings. The reflected signals from buildings increase the travel time of the signals leading to incorrect distance calculation.

A particular interest of this thesis is the use of GPS in Android and the Location API (Application Programming Interface) provided by the open-source mobile operating system. The API provides the app with geographic latitude, longitude, altitude and accuracy from GPS. Accuracy in Android is defined as radius of 68% confidence. In other words, if the location errors are normally distributed, accuracy is the radius of one standard deviation but errors like this do not follow normal distribution. Accuracy is highly dependent upon the number of satellites visible, signal strength, or interruption by external factors. In an open area, GPS on Nexus 4 running Android 4.4.4 measured location with an accuracy of 9–11 meters. Having high rise buildings or trees reduced it to 50 meters.

2.4 Motion Sensors

Motion sensors detect motion and output some predefined motion parameter. Almost all the Android smartphones today are equipped with Accelerometer, and Gyroscope. Motion sensors measure on the device's local x, y and z axes (Figure 2.2). Accelerometer measures the acceleration experienced by the phone in m/sec^2. When the phone is standstill, on a table with the screen facing upwards, the x and y axes acceleration reads close to zero and the one in the z axis, which reads acceleration due to gravity is close to $9.8m/sec^2$. Gyroscope measures rate of rotation, in other words, angular speed, of the phone along x, y and z axes.

Our interest lies in gyroscope which measures angular velocity in radians/sec. Due to inherent nature of the sensor, gyroscope measurements tend to drift overtime. Android filters the drift before providing it to the app by using the data from other sensors. The techniques used to achieve these

results are called Sensor Fusion algorithms [15]. Android also provides uncalibrated readings if required by any app.

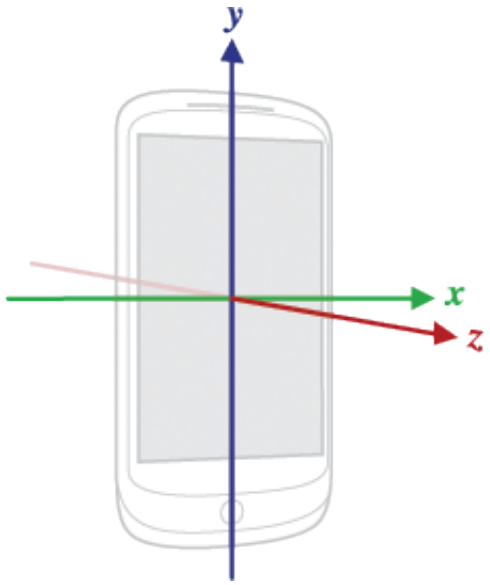


Figure 2.2: Android Device Local Axes [1]

While reading or calculating on any motion sensor data, different types of errors can occur. Except the issue of time delays, all errors are generic and occur in all sensors.

- *Human Error*: These are due to unintentional measurement by human beings. For instance, a person may read a wrong value from the sensor.
- *Systematic Error*: These are because of the system in place and affect the accuracy of the measured value. These are constant offset from the true value. For instance, taking the measurement of magnetometer with a magnet nearby will offset the actual reading by a constant amount.
- *Noise* - It is the random fluctuation in measurement. For instance, even slight perturbation induced by phone speakers, can induce random measurements in gyroscope.
- *Drift*: It occurs when the measurement wanders off the real-world value after some long amount of time.

[1]Image property of Google Inc., under Creative Commons 2.5, `http://developer.android.com/guide/topics/sensors/sensors_overview.html`

- *Offset or Bias*: This error is inbuilt into sensor. It is different from Systematic Error as it is without the influence of any external entity. If the output of a sensor is not zero when the measured property is zero, the sensor has an offset or bias. These values generally need to be subtracted from all the readings.

- *Time Delays and Dropped Data*: Receiving and processing measured data within a limited time period is a quality of Real-Time Operating System (RTOS). Systems, such as Android, are not RTOS. The values reported at any instance may not be the ones being sensed by the sensor. In such a system, data is sometimes delayed or even dropped, resulting in erroneous timestamps.

- *Integration Error*: Many times value measured from a sensor needs to be integrated to find related value. For instance, gyroscope values need to be integrated to find the total angle made by a car in a turn. However, the values in the sensor drift overtime and the error due to zero offset lead to incorrect integration of values. It is practical to subtract the zero offset from the values and perform integration only over a small time interval.

Chapter 3

Threat

In this thesis, we target two threats to the system. The first threat is of privacy invasiveness. Any information that helps adversary determine any user's personal activity is privacy invasive. Proposed VMT tax methodologies collect location information and send it to the server, allowing untrusted servers to access the private information which can be used to infer a lot of other information about the user. The second threat to the system is that of the security of the computation performed inside the vehicle and the integrity of the data received from GPS and odometer. Current generation of cars is vulnerable to many security attacks. It is even possible to change the code running in the car by exploiting the components. Integrity of data received from GPS and odometer is also at risk as the signals received by the system can be modified or spoofed such that it results in less tax.

3.1 Privacy

A VMT tax system that collects location information can be privacy invasive in many forms. A system monitoring the location collects timestamped, fine grained geographical coordinates. The worst case occurs if all the geographic coordinates are sent as it is to the untrusted server for processing and tax computation. Along with computing the required data, the server can infer the movement patterns of the user, vacations taken, restaurants, and hospitals visited, and many more. The users can be personally identified solely based on the location and web search [3]. Home and work locations of individuals can also be identified based on just location traces [4].

The server operator might have good intentions and the user might trust it but the security of the data stored at the server cannot be guaranteed. To overcome it, a VMT tax system may reduce the granularity of the locations

for e.g. by just sending the miles driven in each jurisdiction. This move will reduce adversary's capability to determine some personal activities but depending on the size of the region, it may still leak private information about the whereabouts of a car. Regions with smaller geographic area will allow adversary to predict information with better accuracy.

A 2010 document unearthed by American Civil Liberties Union of North Carolina via a Freedom of Information Act claim reveals that most of the cellular phone service providers store cell towers used by phones for more than a year [16]. There is a very good chance that information about cell towers used by cars equipped with cellular connections capabilities are also stored for more than a year. The cell phone towers cover a large geographical area and it is hard to pin point the location of the device through this data. They can provide a rough estimate of a phone's or car's location. If the phone or car is moving, it can shorten the possible list of locations by the sequence of cell phone towers accessed by the car, the duration of time it was with a cell tower and the maps of all the roads in the area. In this scenario, the user has to trust private location data to cellphone companies. The user doesn't have any other option except to not use the cell phone service.

3.2 Security

A user can tamper with the system installed in the car in order to decrease the tax amount. Tampering can be involved at any stage of the process. A user can modify the signals coming from the wheels such that fewer miles get reported in odometer. He can compromise the electronic system of the car and modify the software or the stored data such that it lowers the tax amount. Depending on the complexity of the process to compromise the electronic system of the car, it may become popular among many users. If the process requires little manual effort, is inexpensive, and is easy to distribute while being hard to detect, we may see many people pursuing it. A difficult process will require a technically skilled or highly motivation person, limiting its spread.

For location, GPS is the most popular choice but it cannot be relied upon for correct location information. GPS signals can be jammed to prevent location recording at all. Having the receiver in a Faraday cage prevents

the signal from reaching the receiver. The user can also use commercial GPS jammers which interfere with the regular signals to render them useless by inducing noise. Fortunately, jamming signals are easily detectable [17]. Spoofing GPS signals to deceive GPS receiver into calculating a fake location is also possible [18]. Considerable work has been done to detect and prevent spoofing but none of the published papers offer a fool-proof technique. A more advanced adversary can always counter the anti-spoof techniques. GPS spoofing has been discussed in section 7.4.

Current odometers are prone to tampering and hence, not reliable if used along for mileage reporting[19]. There are two types of odometer tampering: one changes the stored accumulated miles in the car and the other delivers incorrect data from the wheel while the car is being driven. Most common way to revert mechanical odometers values is to attach the odometer cable to a drill and run it in reverse. Driving a car in reverse also decreases mechanical odometer's value. Reversing in some car models with digital odometers is as easy as attaching a device to OBD-II (On-Board Diagnostics) port of the car and overwriting a new value. Smelecom is one such company selling digital mileage correction equipment [20]. For the other type of tampering, illegal devices are sold in the markets which intercept the signal coming from the wheel and the component responsible for odometer calculation. These devices are man-in-the-middle and modify the values before it reaches the component.

Chapter 4

Requirements

This section aims to enumerate and describe the system requirements to be fulfilled by any proposed solution of VMT tax. The requirements have been modeled on the entities participating in the system. There are four requirement categories - functional, performance, policy, security and 5 participating entities - Tax Authority, User, Car, and Governments.

4.1 Functional Requirements

These requirements define what the system is required to do and specify the minimal abilities that the system must possess. The proposed solution to the problem must be able to perform to meet at the very least of these requirements.

- *The system must tax the User based on the number of miles driven by the Car.*

  The final tax amount charged to the User must be based on the number of miles traveled by the Car. The rate per mile may not be constant and depend on various factors viz. the type/make of the Car or the region in which the Car is being driven.

- *The system must be able to calculate and provide Tax Authority the total tax generated by each Car. The privacy of User must be preserved in the process.*

- *The system must able to provide Tax Authority just enough information to determine the tax share of each Government from the total tax collected.*

  Tax Authority allocates the tax collected to Federal and Regional Governments. The information required to determine the tax proportion

of each region must be conveyed to the Tax Authority. In the whole process, no private data, in form comprehensible to either authority, must be sent.

- *The system must provide Tax Authority the minimum data required to maintain tax records of the User. It may also provide additional data if it does not invade User's privacy.*

  Tax Authority having the minimal information is an ideal condition but it may not be pragmatic. For example, at some point in time, if an administrative authority needs to verify User's identity or the Car, it will require more than the bare minimal information. However, irrespective of the type of information the authority can hold, it must not possess privacy invasive data in a comprehensible manner.

- *The system must provide User with capability to challenge the final bill amount. The data should also be auditable by an independent party who cannot be influenced either by the Tax Authority or the User.*

  The system must allow User to verify and challenge its final tax amount. The system must be capable of storing sufficient data to re-compute the final tax amount. The integrity and authenticity of the data being stored must by verifiable. The system may support tax computation by an independent third party to resolve the conflict.

4.2 Performance Requirements

- *The number of miles traveled must be collected from a reliable source. The source may be the odometer of the Car.*

  As mentioned by Hanley and Kuhl, GPS is not a reliable technology for continuous location monitoring [21]. The signals are satellite dependent and any disruption in the signal (for e.g. by high rise buildings, tunnels, dense trees, etc.) makes location determination unreliable and hence the calculation of number of miles driven. Odometer on the other hand can determine the number of miles travelled reliably with very less error. GPS can be used to verify the number of miles travelled with some error.

- *The Car must be able to send required information reliably to Governmental Authority at specified intervals.*

  The Car will gather information on the number of miles travelled in a particular region. This information will need to be communicated to the Governmental Authority. The data transfer interval needs to be at least as frequent as Government Authority specifies it. Accordingly, the communication mechanism and the medium used for transmission must be reliable enough to support the delayed reporting at appropriate batch times.

- *The communication mechanism between the Car and the User must have adequate reliability to guarantee the deliverability of the information. The mechanism should also be user-friendly to the ordinary people.*

  Conditions of appropriate time intervals and reliable communication as mentioned above apply here. Besides, the system should ensure that the communication mechanism is easy to use for ordinary people. It should also account for cases when a lay man is unable to fulfill its duties.

- *The system must ensure that the data made available/sent to the user must have a copy. It must also ensure that the User privacy is not violated in the process.*

  If some data is sent to the user, there is a good chance that the User might lose that particular piece of data. To deal with such scenarios, the system needs to make the data redundant. The data should also be reliably retrievable whenever the requirement arises.

- *The system must support the required resolution of the data.*

  Granularity of the data (odometer and GPS) provided by the hardware must be equal or more than the one specified by Tax Authority. The software of the system must be able to function smoothly for the designed granularity. If the granularity is high, the size of the data to be processed and its processing time will rise up for the same inexpensive hardware. The hardware required to process more granular data in sufficient time may cost extra.

- *The system must ensure the collection data rate as specified by Tax Authority.*

  The data collection rate must provide smooth functioning. For example, the GPS and the odometer data might be recorded every 10 seconds. If the Car enters another region within that 10 second period, the collected data could be interpolated to calculate the miles in each region. The collection rate will depend on the acceptable error specified by Tax Authority.

4.3 Security Requirements

These requirements provide the security and privacy specification for the system.

- *The system must ensure confidentiality, integrity and availability of all the data transmitted from one entity to another.*

  All the data transmitted between any two entities in the system must be encrypted to provide confidentiality and integrity verification. A person eavesdropping on the data must not be able to view any part it or be able to modify it without getting detected. The system must also ensure the availability of data to be transmitted.

- *The system must provide authentication for all the senders.*

  Whenever a data is being sent from one entity to another, the receiver must be able to authenticate the sender of the data. For instance, if the Car is transmitting data to Tax Authority, the authority should be able to verify that the data was indeed sent by the Car.

- *The system must ensure to the best of its ability that the hardware and software of the device used for data collection and computation is tamper resistant.*

  Any adversary must not be able to tamper with the hardware and the software of the device being used for data collection or tax computation. If one does tamper with any part of the device, the device must be able to detect it and take appropriate actions. The device should

notify the Tax Authority or the entity responsible. The insecurity of the electronics components of a car have been shown in [11, 5]. The authors were able to reverse engineer and compromise many Electronic Controller Units (ECUs). Much critical functionality like odometer reading and brakes were compromised along with many others.

- *The system must be able to detect any tampering with the hardware or software. In case of any suspicious activity, it should be able to take necessary steps to prevent any damage.*

  A malicious user might try to tamper with the device collecting the data or provide fake data to the device. The device must use other mechanisms to determine the possibility of tampering. Various methods to detect and protect against the device tampering have been discussed in [22]. For example, an adversary may block GPS signals all together using Faraday's cage. The system should be able to detect it and take necessary steps. In this case, reading odometer data or installing an accelerometer in the device to detect car's movement can help verify if the car is moving. If the car is moving and GPS does not detect any signal, the tax calculation may be done at some pre-defined high rate.

- *The system must ensure that Tax Authority does not receive User's fine location data in readable form. If it receives it in encrypted form, the system should ensure that there are no information leaks from the data. If possible, the system can also ensure that the location data is not available to anyone except the User.*

- *The system should minimize the side-channel and covert channel information leaks.*

  All the known side-channel attacks against the devices deployed in the system must be studied. They should be rendered fool-proof for as many attacks as possible. The hardware and software may be researched for any side-channel or covert channel information leak.

4.4 Policy Requirements

- *The system must incorporate policy on information loss from the Car.*

The system must include policies to tackle the situations resulting in data loss from the Car. It might occur due to an adversary attempting to tamper with the device or uncontrollable circumstances like car crashes and storage destruction.

- *The system should support uniform policy across a region. If the prices are dynamic, the User must be notified about the prices through some mechanism.*

  The policy of uniform rate in a region can be compared with the current system of gas stations. When a person visits a gas station, he makes a decision of buying gas based on the price at the station. If he feels that the price is too high, he can try another gas station. He has a choice of selecting the gas station based on the pricing. In our case, if the rates vary within a region, the User must be made aware of the rates for driving on a road. He must be able to take another path, if available, if rates seem unfavorable to him.

- *The system must include a mechanism to check sensors and computation for correctness.* The system must support a verification mechanism to determine if the hardware and software are working as expected. There might be a human element involved in verification process. Periodic random checks on instruments can demoralize tampering.

- *The system should have support for scenarios where the driver is not the owner of the car and is not responsible for tax payment of the car.*

  Rental cars are the perfect example of this scenario. The system must develop some mechanism to charge users, either though rental company or directly to the driver of the car. It must be able to preserve privacy of the driver in such cases too.

4.5 Participating Entities

These entities directly affect the working of the system. Later in the section, we recognize some stakeholders in the process who either affect the system or get affected by it in some manner.

4.5.1 Tax Authority

This is the central governing body for the mileage-based taxation. It is a part of federal government but can also be a private entity on a contract.

Functions

- The authority, with the help of other government agencies, specifies the region boundaries and sets the price per mile of each region.
- It specifies the granularity of the data to be used for calculating the taxes. It is presumed that the granularity defined by the Governmental Authority is practical and can be achieved through inexpensive hardware.
- It specifies the minimum frequency for data transmission from the Car. The time interval must be practical enough to allow for system's smooth functioning
- It must also specify the maximum acceptable level of errors in the collected raw data and tax calculations
- It must decide on the life of the data. If the car is storing data locally, it needs to specify the minimum time for which the data must be kept in the car. It also should specify the maximum time.
- It must define reliable communication and make policies for unreliable communication.
- The entity must develop rules and regulations for rental vehicles and taxis.

Responsibilities

- It issues tax statements to the User, collects taxes and maintaining tax records.
- It is responsible for allocating the collected taxes to the transportation authorities of respective regional governments and federal government.

The amount allocated should depend on the number of miles driven by the cars in the region.

- It may be responsible for the mechanism to compute taxes. The work can be taken by a separate entity.
- It is responsible for handling User's complaints and must provide the services to process the audit requests.

4.5.2 User

A person registered with the Tax Authority for paying the mileage-based tax. All the tax statements are issued in his/her name and he/she is the one responsible for paying the tax for an automobile. User may be the owner of the Car. It is up to the Tax Authority to decide the rules.

4.5.3 Car

It is the vehicle for which the tax is charged to the User. Number of miles driven by the Car along with the type and model of the car should determine the amount of tax.

4.5.4 Governments

These include the Federal government as well as the Regional Governments administering tax jurisdictions. These entities handle the transportation departments' revenues and finance. They may be involved in developing or maintaining surface transportation infrastructure.

4.5.5 Stakeholders

Stakeholders are the entities who affect or can be affected by the system. The affects may be monetary, political, contractual, technological, etc. Stakeholders, by definition include participating entities.

Automobile Manufacturer

Implementing the system on the automobiles may affect the way cars are manufactured, programmed or serviced.

- All cars after 1996 have standard interfaces (OBD-II) to read the data from the car. The electronic architecture is similar in most of the cars (AUTOSAR) which helps in collection of data, such as from odometer without requiring support from car manufacturer. Some cars have different systems which require cooperation from the car manufacturer.
- If the system gets mandated, car manufacturers might want to integrate the whole system in the car itself at the time of manufacturing.
- Car manufacturers will have to provide a methodology to verify that the systems in the car are working correctly. It will be required when a User wants to audit the tax imposed and the auditor wants to verify the correct functioning of the equipment.

Communication Service Providers

They provide the service to transfer data to and from the Car to other entities. They will have contract with the User, the Car manufacturer or Tax Authority for reliable communication service.

Electronic Equipment Providers

These are the companies providing the equipment for mileage calculation, cryptographic computations, etc. Their products and business will affect and be affected by the system. These companies will be responsible for providing secure and certified equipment with reliable performance and robustness. They may be responsible for equipment management and service.

Standardizing Organizations

Organizations, such as IEEE (Institute of Electrical and Electronics Engineers) and SAE (Society of Automotive Engineers), who standardize various mechanisms and protocols, might be involved in providing suggestions for

standardizing some solution for VMT tax collection system. The process will involve standardizing the hardware used for this purpose, the communication used to transfer the information, methodology used for tax calculation, computation methodologies, tax distribution system, and many more.

Law Enforcement and Lawmakers

Agencies responsible for maintaining and upholding law will get additional responsibilities. Lawmakers will be enacting laws for the agencies functioning and policies for tampering with the system. The will also set up rules to resolve disagreements between any two entities.

Law Firms

This system will require many contractual agreements among various entities. Law firms will be involved in detailed contract writing and representing entities in legal proceedings for resolving conflicts.

Civil Rights

The system involves user privacy component and will affect the groups vouching for civil rights.

Energy Providers

The model of charging price per mile will change the number of miles driven by users. It will affect the energy consumption, directly affecting the supply and revenue of the energy providers, including but not limited to electric utility companies and gas companies.

Traffic Analysis

Different price per mile for regions will affect the amount of traffic on different routes. Traffic flow analysts will have to incorporate tax rate in their calculation even for regular roads and recalculate the flow. For future projects,

the rate might be an important role and might be varied to change traffic flows and develop infrastructure accordingly.

Chapter 5

Design

The design of PPVMT is based on the system requirements presented in section 4. While the design incorporates most of the functional and security requirements, we do not focus on policy requirements. The performance requirements have been discussed more in section 6.

The design of PPVMT primarily targets two problems: location privacy preservation and tamper resistance. User's location information should not be accessible by anyone except the user and the user should not able to tamper with the system easily. The design doesn't aim to make the system tamper proof. It aims to increase the cost and required effort to tamper with the system by detecting tampering, imposing fines and making it hard to tamper.

5.1 Entities

These are similar to the entities in section 4.5. One extra entity has been added and more functionality and requirements have been added to the existing entities to support the design.

Car

It is the vehicle for which the tax is charged to the User. The car uses GPS, inertial motion sensors (accelerometer and gyroscope) and odometer. The Car also has computation power to perform cryptographic operations and capability to generate pseudo-random bits required for secure computations.

Currently, all cars some with odometers but only few cars models come with inbuilt GPS functionality. Gyroscopes are not present in current cars but advancements in MEMS (MicroElectroMechanical Systems) technology

have brought down the cost of motion sensors significantly. For instance, a 3-axis accelerometer and a 3-axis gyroscope are now available for less than $5.

User

The person registered with the Tax Authority for paying the VMT tax. All the tax statements are issued in his name and he is the one responsible for paying the tax for an automobile.

Independent Server

There are at least two of these and the design requires that these servers are non-colluding, preferably administered by different authorities who are bound by law against collusion. Diverse arrangements are possible for the entities handling the servers - the servers can be under private entities, one under a private entity, and one under government, or both the servers under the government. Having Independent Servers under private entities may be perceived positively among general populace [21].

Tax Authority Server

The Tax Authority Server is responsible for determining the tax of Users based on their mileage in each geographical region and allocating the total tax collected to Regional and Federal Governments.

Governments

These are the government agencies responsible for surface transportation infrastructure in their administrative regions.

Key Management Server (KMS)

This server stores and certifies the public certificate of all the entities in the system. To verify the integrity and authenticity of the communication, all

entities contact the Key Management Server to get the sender entity's public certificate.

5.2 Architectural Framework

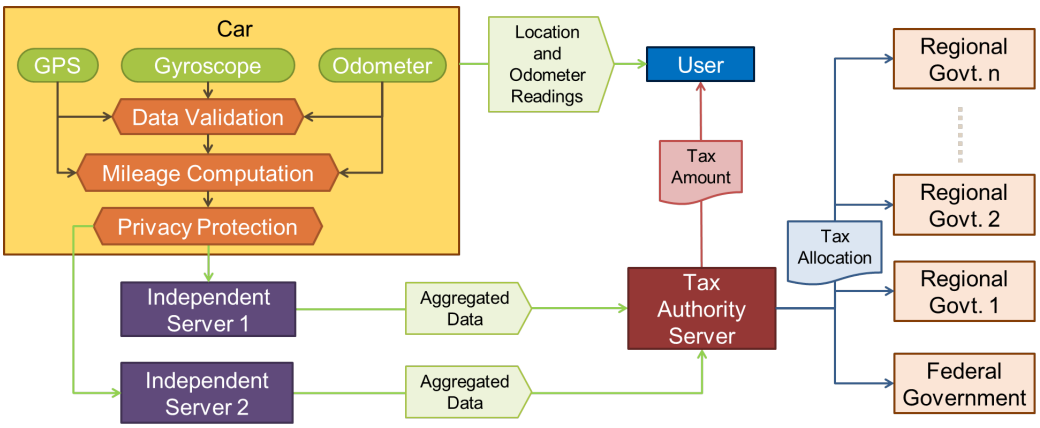


Figure 5.1: Architectural Framework

The system ensures that location information of any user is not leaked to any entity, yet the Tax Authority is able to compute taxes. The authority does not learn anything except the tax of each user, and the tax share of the federal and each regional government. The system also provides tamper resistance against odometer and GPS spoofing. The design allows a copy of the data generated from the car (fine grained location, odometer values, and computed data) to be provided to the user, allowing him to do his own calculation and verify the tax amount.

Figure 5.1 present the design of Privacy-Preserving Vehicle Miles Traveled (PPVMT) tax system. The Car receives the location coordinates from the GPS unit. The Car stores the regional boundaries and the pre-determined tax rates of each region. The software running in the Car determines the region in which the car is being driven and records the miles driven through odometer data. Based on the policies and hardware, the system stores the values for the specified period of time. When it is time to send the mileage data to the server, it performs the privacy preserving cryptographic operations (Section 5.4) and sends the data to Independent Servers. The region boundaries are stored on the Car itself and not queried from the server as es-

tablishing a network connection to the server can give away the approximate location of the car.

To lower his tax amount, a user can compromise the computation, spoof the GPS signals or modify the data coming from odometer to report less mileage. The design detects GPS signal spoofing and odometer tampering by inertial motions sensors. This framework does not detect or prevent software tampering. It assumes that all the computations in the car are performed in the manner as expected by an untampered software. The car runs a validation algorithm on the stream of data coming from GPS, accelerometer, gyroscope and odometer which verifies the integrity of the data by cross-checking the values. This can be done in many different ways, some of which has been described in section 5.5. This design concerns itself only with fraud detection. Reaction of the system to the fraud, for instance, whether to charge a higher rate, report it to a server, mention the location and the time of the car, is dependent on the law and policy decided by the Tax Authority.

This design consists of 2 Independent Servers. Each car splits its miles into two values and sends it to each of the Independent Servers. This helps in preserving privacy as none of the servers can find out the number of miles traveled by any car in any of the regions. The servers perform aggregation on the collected data periodically and send the aggregated value to the Tax Authority server.

The Tax Authority server aggregates the data received from all the Independent Servers, calculating tax owed by each individual User and computing the amount to be allocated to each regional government. Based on the policies the Tax Authority notifies the User of the tax owed.

The User gets all the data signed by the Car. It includes the data that is sent to the Independent Server as well as the private timestamped location data. The location data enables the user to calculate its own tax as the prices of each region are public. If there is any discrepancy in the amount, the user can present the data signed by the car to the Tax Authority to get the issue rectified. The User can get the data from the Car in multiple manners. The Car either uses the inbuilt Internet connectivity to send the data to the User, or transfers the data to the smartphone through wireless Bluetooth or by wired connection, such as USB. It can also transfer the data into a storage device which can be plugged into the Car. These methods are a matter of availability, technical capability, policies, and user-friendliness.

The frequency with which the data needs to be collected by the Independent Servers and the Tax Authority server is determined by the Tax Authority. The system can support highly periodic data reporting as well as with longer intervals without compromising the location privacy of the User.

All network communications among entities use Secure Sockets Layer (SSL) to prevent any eavesdropping or data modification. There exists a trusted Certificate Authority (CA) which signs the public-key of every entity. If the private key of a server gets leaked, it's easy to replace but for cars, replacing the key may not be easy. If the Car has Internet connectivity enabled, the key can be updated over the network else the Car might need to be taken to a certified mechanic.

5.3 Tax Computation

The methodology to compute total tax for each individual user as well as each region has been shown in table 5.1. The technique has been described using three cars and three regions but as we will see, the technique can be scaled to any number of cars and any number of regions. The three cars are - CAR_1, CAR_2, CAR_3, and the three regions are - RGN_A, RGN_B, and RGN_C. The price per kilometer of each of the region is $\$P_A$, $\$P_B$ and $\$P_C$. M_{jk} represent the total number of KM driven in region j by car k.

Total tax owed by the user of the Car is the sum of tax owed in each region which in turn is equal to price per KM times the distance driven. For car CAR_k, the tax owed will be $P_A * M_{Ak} + P_B * M_{Bk} + P_C * M_{Ck}$, as shown in the last column of table 5.1.

Similarly, tax for each region is the total distance driven by all the vehicles in that region times the price per KM of the region. For RGN_j, the tax amount from three cars comes out to be $P_j * (M_{j1} + M_{j2} + M_{j3})$, as presented in the last row of table 5.1.

Generalizing the solution for N cars and L regions,

$$\text{Tax owed by } CAR_k = \sum_{i=1}^{L} P_i * M_{ik} \tag{5.1}$$

Table 5.1: PPVMT Tax Computation

Car\Region	RGN_A ($\$P_A$ / mile)	RGN_B ($\$P_B$ / mile)	RGN_C ($\$P_C$ / mile)	Tax of each Car
CAR_1	M_{A1}	M_{B1}	M_{C1}	$P_A*M_{A1} + P_B*M_{B1} + P_C*M_{C1}$
CAR_2	M_{A2}	M_{B2}	M_{C2}	$P_A*M_{A2} + P_B*M_{B2} + P_C*M_{C2}$
CAR_3	M_{A3}	M_{B3}	M_{C3}	$P_A*M_{A3} + P_B*M_{B3} + P_C*M_{C3}$
Tax of each Region	$P_A*(M_{A1}+M_{A2}+M_{A3})$	$P_B*(M_{B1}+M_{B2}+M_{B3})$	$P_C*(M_{C1}+M_{C2}+M_{C3})$	$P_A*(M_{A1}+M_{A2}+M_{A3}) + P_B*(M_{B1}+M_{B2}+M_{B3}) + P_C*(M_{C1}+M_{C2}+M_{C3})$

and

$$\text{Tax amount for } RGN_j = P_j * \sum_{i=1}^{N} M_{ji} \tag{5.2}$$

5.4 Preserving Location Privacy

As mentioned in section 3.1, sending the fine-grained geographic coordinates as well as send the total miles covered in each region is privacy invasive. To protect the privacy, we propose a mechanism based on additive secret sharing scheme. A similar technique has been proposed in the past for aggregating smart-meter data in a privacy enhanced manner [23].

For each region, the Car generates a random integer r and subtracts it to the miles driven in that region to yield r'. r_{A1} and r'_{A1} represent the integers generated by CAR_1 for RGN_A such that

$$r_{A1} + r'_{A1} = M_{A1} \tag{5.3}$$

Table 5.2 shows all the integers generated for the three cars and the three regions and the aggregation after replacing the miles with r and r'.

If the miles are not in integer, they can be multiplied by a power of 10 to make it an integer. For instance, if the miles reported extend two places after the decimal, multiplication by 100 will yield the required integer. The final aggregated sum can be offset by the same power of 10 to find the actual value.

A car sends all its random numbers r to Independent Server 1 and r' to

Table 5.2: Tax computation with secrets

Car\Region	RGN_A ($\$P_A$ / mile)		RGN_B ($\$P_B$ / mile)		RGN_C ($\$P_C$ / mile)		Tax of each Car
	IS 1	IS 2	IS 1	IS 2	IS 1	IS 2	
CAR_1	r_{A1}	r'_{A1}	r_{B1}	r'_{B1}	r_{C1}	r'_{C1}	$P_A*(r_{A1}+r'_{A1}) + P_B*(r_{B1}+r'_{B1}) + P_C*(r_{C1}+r'_{C1})$
CAR_2	r_{A2}	r'_{A2}	r_{B2}	r'_{B2}	r_{C2}	r'_{C2}	$P_A*(r_{A2}+r'_{A2}) + P_B*(r_{B2}+r'_{B2}) + P_C*(r_{C2}+r'_{C2})$
CAR_3	r_{A3}	r'_{A3}	r_{B3}	r'_{B3}	r_{C3}	r'_{C3}	$P_A*(r_{A3}+r'_{A3}) + P_B*(r_{B3}+r'_{B3}) + P_C*(r_{C3}+r'_{C3})$
Tax of each Region	$P_A*(r_{A1}+r'_{A1} +r_{A2}+r'_{A2}+r_{A3}+r'_{A3})$		$P_B*(r_{B1}+r'_{B1} +r_{B2}+r'_{B2}+r_{B3}+r'_{B3})$		$P_C*(r_{C1}+r'_{C1} +r_{C2}+r'_{C2}+r_{C3}+r'_{C3})$		$P_A*(M_{A1}+M_{A2}+M_{A3}) + P_B*(M_{B1}+M_{B2}+M_{B3}) + P_C*(M_{C1}+M_{C2}+M_{C3})$

Table 5.3: Aggregation on Independent Server 1

Car\Region	RGN_A ($\$P_A$ / mile)		RGN_B ($\$P_B$ / mile)		RGN_C ($\$P_C$ / mile)		Tax of each Car
	IS 1	IS 2	IS 1	IS 2	IS 1	IS 2	
CAR_1	r_{A1}		r_{B1}		r_{C1}		$P_A*r_{A1}+P_B*r_{B1}+P_C*r_{C1}$
CAR_2	r_{A2}		r_{B2}		r_{C2}		$P_A*r_{A2}+P_B*r_{B2}+P_C*r_{C2}$
CAR_3	r_{A3}		r_{B3}		r_{C3}		$P_A*r_{A3}+P_B*r_{B3}+P_C*r_{C3}$
Tax of each Region	$P_A*(r_{A1}+r_{A2}+r_{A3})$		$P_B*(r_{B1}+r_{B2}+r_{B3})$		$P_C*(r_{C1}+r_{C2}+r_{C3})$		*Data sent to the Tax Authority*

Independent Server 2 (Table 5.3 and 5.4). Until and unless the two servers collude, all the numbers will look completely random to either of the servers. The servers cannot figure out the number of miles traveled by the car in any of the regions. Both the servers aggregate the random integers in the manner similar to finding tax for each user and region. It has been shown in the last column of table 5.3 and 5.4 for the sample case. Generalizing for N cars and L regions, computation on Independent Server 1 will be

$$\text{Aggregation for } CAR_k = \sum_{i=1}^{L} P_i * r_{ik} \tag{5.4}$$

and

$$\text{Aggregation for } RGN_j = P_j * \sum_{i=1}^{N} r_{ji} \tag{5.5}$$

Independent Server 2 will perform exactly the same aggregation on r'.

Table 5.4: Aggregation on Independent Server 2

Car\Region	RGN_A ($\$P_A$ / mile)		RGN_B ($\$P_B$ / mile)		RGN_C ($\$P_C$ / mile)		Tax of each Car
	IS 1	IS 2	IS 1	IS 2	IS 1	IS 2	
CAR_1		r'_{A1}		r'_{B1}		r'_{C1}	$P_A*r'_{A1}+P_B*r'_{B1}+P_C*r'_{C1}$
CAR_2		r'_{A2}		r'_{B2}		r'_{C2}	$P_A*r'_{A2}+P_B*r'_{B2}+P_C*r'_{C2}$
CAR_3		r'_{A3}		r'_{B3}		r'_{C3}	$P_A*r'_{A3}+P_B*r'_{B3}+P_C*r'_{C3}$
Tax of each Region	$P_A*(r'_{A1}+r'_{A2}+r'_{A3})$		$P_B*(r'_{B1}+r'_{B2}+r'_{B3})$		$P_C*(r'_{C1}+r'_{C2}+r'_{C3})$		Data sent to the Tax Authority

Table 5.5: Aggregation on Tax Authority Server

Car\Region	RGN_A ($\$P_A$ / mile)		RGN_B ($\$P_B$ / mile)		RGN_C ($\$P_C$ / mile)		Tax of each Car
	IS 1	IS 2	IS 1	IS 2	IS 1	IS 2	
CAR_1	r_{A1}	r'_{A1}	r_{B1}	r'_{B1}	r_{C1}	r'_{C1}	$[P_A*r_{A1}+P_B*r_{B1}+P_C*r_{C1}] + [P_A*r'_{A1}+P_B*r'_{B1}+P_C*r'_{C1}]$
CAR_2	r_{A2}	r'_{A2}	r_{B2}	r'_{B2}	r_{C2}	r'_{C2}	$[P_A*r_{A2}+P_B*r_{B2}+P_C*r_{C2}] + [P_A*r'_{A2}+P_B*r'_{B2}+P_C*r'_{C2}]$
CAR_3	r_{A3}	r'_{A3}	r_{B3}	r'_{B3}	r_{C3}	r'_{C3}	$[P_A*r_{A3}+P_B*r_{B3}+P_C*r_{C3}] + [P_A*r'_{A3}+P_B*r'_{B3}+P_C*r'_{C3}]$
Tax of each Region	$P_A*(r_{A1}+r_{A2}+r_{A3}) + P_A*(r'_{A1}+r'_{A2}+r'_{A3})$		$P_B*(r_{B1}+r_{B2}+r_{B3}) + P_B*(r'_{B1}+r'_{B2}+r'_{B3})$		$P_C*(r_{C1}+r_{C2}+r_{C3}) + P_C*(r'_{C1}+r'_{C2}+r'_{C3})$		Computation on Tax Authority Server

Once all the operations for this set of data is completed, both the servers send the aggregated values for each user and each region to Tax Authority Server. In table 5.3 and 5.4, it corresponds to the last column and the last row. Tax Authority server aggregates the corresponding value from each of the Independent Server to yield the final value as shown in table 5.5. The value of each element in last row and last column of table 5.5 is same as table 5.2, which in turn, because of the equation 5.3, is same as table 5.1.

5.5 Tamper Resistance

A malicious user can try to alter the functioning of the system to reduce his tax amount. GPS and odometer are prone to exploitation (Section 3.2) as both of them can be compromised in some manner to report fake locations and miles. To protect the system, tamper-resistant techniques needs to be

applied [24].

Inertial motion sensors do not rely on any external electromagnetic signal but on the mechanical forces acting on the sensors. These sensors are more trustworthy than GPS or odometer as they are usually embedded in the machinery or on the chip. Fiddling with them require high technical skills. Also, they work on complementary set of data. Providing them with fake data will again require high skill.

5.5.1 Accelerometer and Odometer

Accelerometer provides the acceleration of the car. If the data from the odometer has high granularity and can be updated with a low latency (10 times a second should work), then the distance data can be numerically differentiated two times with respect to time to determine the acceleration. This acceleration can be matched with accelerometer at any instance of time to determine if the odometer has been tampered.

5.5.2 Accelerometer and GPS

GPS provides geographical coordinates with some accuracy. For cheap GPS receivers, such as the ones installed in Smartphones, the accuracy is around 9 meters. With errors as huge as 9 meters, it is impossible to detect fine-grained position changes and hence, the acceleration. It might be possible with highly accurate and precise GPS receivers but the number of errors that can crop up in data transmission and receiving is really large, making the cross-checking impractical to consider.

5.5.3 GPS and Odometer

Though GPS and odometer are both untrusted, comparing them can be beneficial to get long term data comparison. The distance covered by the GPS and the odometer over a long period of time should be the same within some error bound. If the distances are very different, it means either one of it has malfunctioned or has been tampered.

5.5.4 GPS and Gyroscope

Gyroscope provides the rotation speed of the sensor. Whenever a vehicle takes a turn, the gyroscope provides the angular speed which can be integrated over short intervals to determine the angle turned by the vehicle. The coordinates provided by the GPS should also conform to the angle turned by the vehicle. Due to inaccuracy, noise and errors in data the angle covered might not always be the same. If turning angles don't match for many turns, it can be concluded that the GPS has been tampered.

The above methodologies don't ensure that all the spoofing and signal modifications will be detected. It ensures that unsophisticated attacks can be sensed. If the attacker modifies the GPS and the odometer values such that it always lies below the error thresholds, the system will not be able to detect it.

Chapter 6

Implementation

This chapter presents the implementation of the design described in the previous chapter. We propose 'Car-as-a-Smartphone' model and develop a proof-of-concept of tamper resistance against GPS signal spoofing. The authenticity of GPS coordinates is established by comparing the data with gyroscope measurements in a running car. We also implement the additive secret sharing technique for privacy protecting and evaluate its performance.

6.1 Car-as-a-Smartphone Model

In automotive industry, companies are not found collaborating as much as computer industry. There is hardly any sharing of code with general public. Automotive companies maintain competitive edge over each other based on their technologies and implementation of existing technologies. Giving it out as open-source hasn't been a part of the industry. Also, giving out the complete design of the cars and source code of the programs installed in cars might reveal security loopholes and vulnerabilities, which may danger the safety and security of the cars running on the road. Even if updates patching the vulnerabilities are delivered, a big percentage of cars do not support web based software update. Applying updates will require taking the car to the mechanic and is at the sole discretion of the customer while costing them money.

In the last decade, there have been positive developments with technology and software sharing. In 2003, leading automotive manufacturers and component suppliers started a worldwide cooperation to develop an open and standardized software architecture known as AUTOSAR (AUTomotive Open System ARchitecture) [25]. It is based on the principle, "Cooperate on standards, compete on implementation". It standardizes basic system

functions and functional interfaces through specifications and automobiles have to fulfill them to be AUTOSAR compliant. GENIVI Alliance is another industry led effort [26]. In their own words, "GENIVI is committed to driving the broad adoption of an In-Vehicle Infotainment (IVI) open-source development platform."

Open Automotive Alliance (OAA), the latest among all, is an alliance of technology companies (including Google) and automotive manufacturers to bring Android to automobiles [27]. It aims for compatible use of Android devices with cars. In the long run, the alliance also plans to enable the car itself as a connected Android device.

Looking at the trends, we propose 'Car-as-a-Smartphone' model. It hypothesizes that the various features available in a near future car will be very close to those of a smartphone. We emulate car functionality present in our design through Android apps. As mentioned in section 2.2, the electronic system of current cars has been shown to be highly vulnerable to devastating attacks. We assume that the electronic system of our Car can neither be overwritten by an unauthorized entity nor its functionality be modified by malicious program or user. The operating system can perform correct computation and the software doing the calculation for tax cannot be modified or removed.

6.2 Information Flow

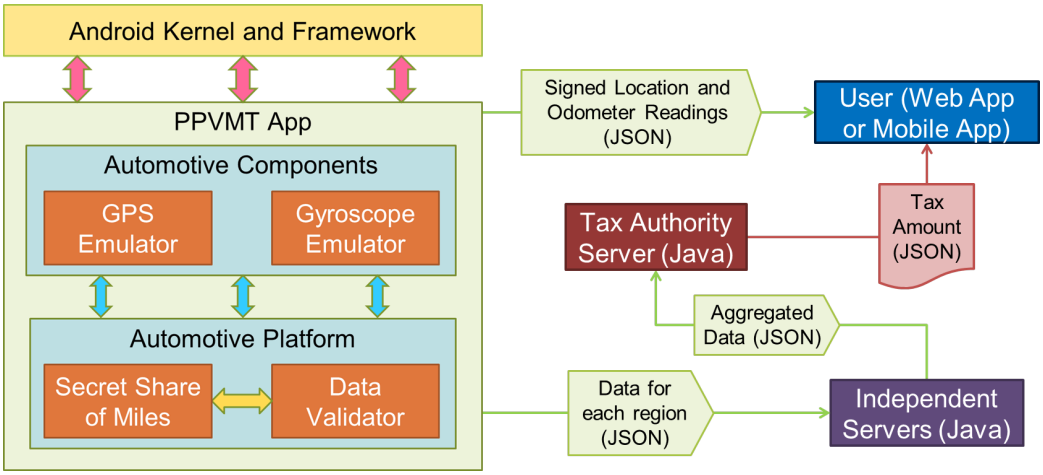


Figure 6.1: System Implementation

Adopting the 'Car-as-a-Smartphone' model, we perform all the operations of a Car on Android Smartphone. The app on the phone collects data from the GPS and the gyroscope, validates the GPS data using the techniques discussed in section 6.4. If the data passes the validation test and is found to be valid, the app calculates the miles traveled in each region and stores it in the phone to be sent to Independent Servers.

According to the policy set by the Tax Authority, when the time arrives to send the data to Independent Servers, the app splits the total miles into random values for each region and sends it to the Independent Server along with an identifying number of the app which in actual deployment would represent the identifier of the car. The app sends all the data serialized as a JSON string which gets deserialized at the servers.

Independent Servers aggregate the data as explained in section 5.4. The time required for aggregation varies with the number of regions and number of cars being aggregated (Section 6.6). Independent Servers send data to the Tax Authority Server once the aggregation is complete. The time of sending depends on the policy set by the Tax Authority and can be varied.

Tax Authority aggregates the data as show in table 5.5. Once the amount has been computed for each user, the server sends the information about the tax amount to the User through a web application in a browser or on a mobile app.

The User gets to download and see the data stored by the PPVMT app. The data consists of timestamped geographic coordinates from GPS. This data is used the User to do his own calculation of the tax and verify the amount calculated by Tax Authority.

6.3 Data Collection

Figure 6.2 plots gyroscope values recorded by an Android phone mounted on the windshield of a car. While recording the data, the phone was mounted in landscape view which led to the phone turning along x axis whenever the car took a turn. These are visible as spikes in the data along x-axis. The data was recorded using PPVMT Data Collector App (details in Section 6.3.1). Android supports retrieval of measured value with different frequency. For tamper resistance, reading the gyroscope value 10 times a sec turned out to

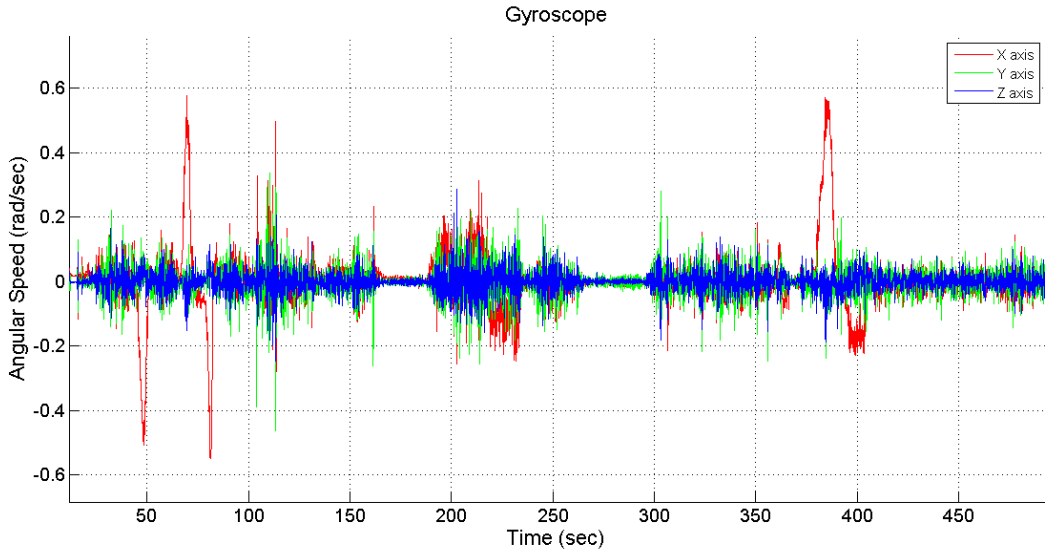


Figure 6.2: Raw Gyroscope Values

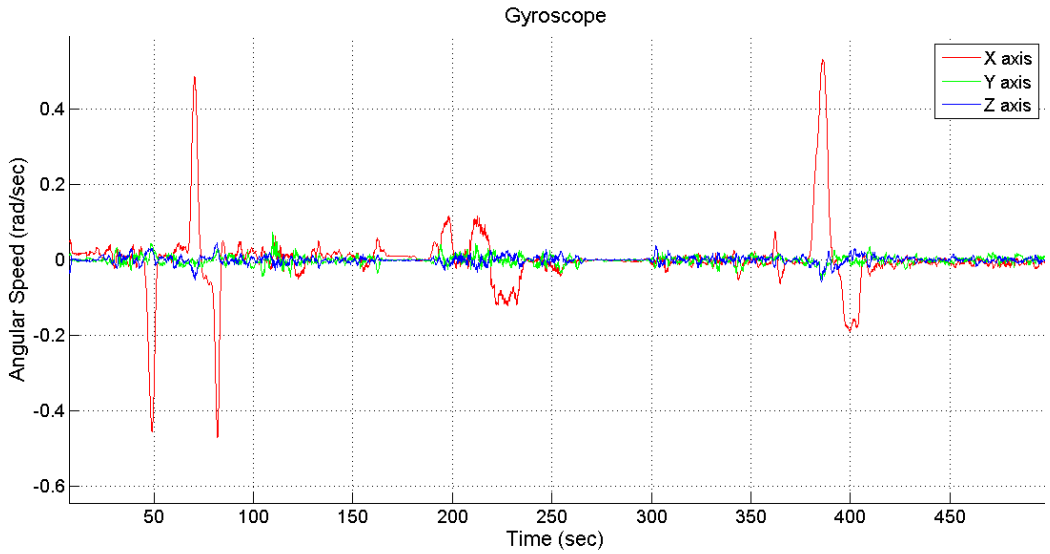


Figure 6.3: Simple Moving Average (SMA) of Gyroscope Values

be sufficient for turn detection and integration to find turn angle.

There is a lot of noise on the data along y and z axes. High spikes for longer periods on y and z axes would mostly occur when a car totals. Spikes along y axis occur when the car breaks, bringing suspension system into action, which rotates the car about y axis. As the car resets, the car rotates back by almost the same amount leading to another spike in opposite direction. Spikes around z axis occur when the car is turning or shifting lanes. The

suspension system rotates the car a small angle along the z-axis, registering reading on the gyroscope. Besides these, random wobble of the device due to imperfect holder, wobbling of the car due to uneven road and machinery also leads to noise.

Averaging the gyroscope readings eliminates the noise as the noise values are equally probable in opposite direction. Figure 6.3 plots the Simple Moving Average (SMA) of the same set of values in figure 6.2. The average at any instance of time is the average of last 1.5 seconds of data. The noise along the y and z axes gets eliminated and only the data along the x-axis, representing the turns remains.

Similar to motion sensors, Android supports location retrieval at user defined frequency but in the experiments, Android did not update GPS location more than once per second. Such low frequency prevents the use of GPS for determining linear acceleration but it is still useful for finding the turn angle.

6.3.1 PPVMT Data Collector App

Figure 6.4 is the main screen of the app which was used to collect data for PPVMT Tamper Resistance. The data collected from the app was used to make data validation model as presented in section 6.4.

For every trip, the app collects timestamped GPS location, gyroscope, and linear acceleration values 10 times per second. The driver of the car entered the miles from the odometer manually. This was to find out if the total distance calculated by GPS is same as the one reported by odometer. For all the 33 trips recorded, total miles from GPS matched the odometer readings. The odometer in the car had the least count of 1 mile.

When the user presses *Start Recording* button on the app, it starts a Service which runs indefinitely until the user presses *End Recording*. The Server spawns a thread upon starting which records the GPS, gyroscope and linear acceleration data 10 times a second. Though it writes the data to a zip file to conserve space and reduce the data, the amount of data generated was considerably small even for long trips. The data sizes and other metrics are discussed in section 6.6.

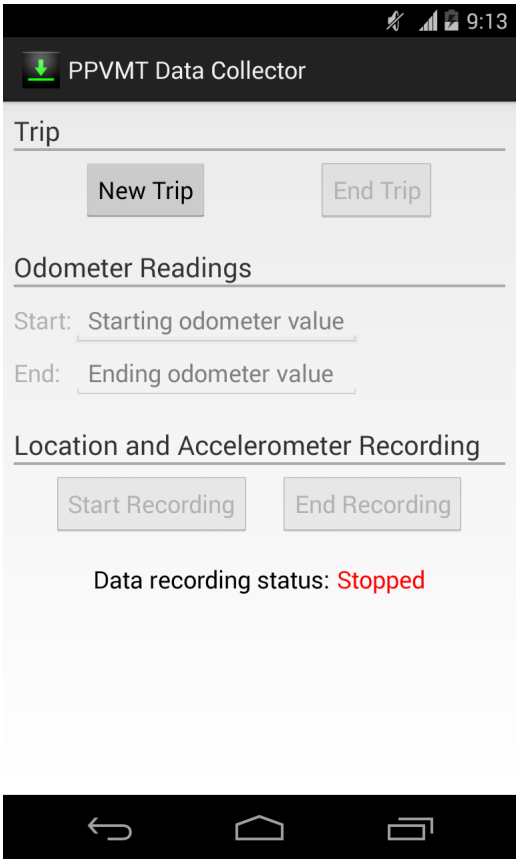


Figure 6.4: PPVMT Data Collector - Main Screen

6.4 Data Validation Algorithm

It aims to validate the GPS coordinates from gyroscope data on an Android smartphone. The algorithm consists of 4 parts:

1. Determining turn from gyroscope measurements
2. Combining multiple turns into one if the they are within some seconds
3. Calculating angle from gyroscope and GPS data
4. Computing the validity score of the turn on various factors

6.4.1 Determining turn from gyroscope measurements

Turn detection is based on threshold filters. A spike in the x-axis can be caused either by a turn, a curve in the road or by lane shifts. Therefore,

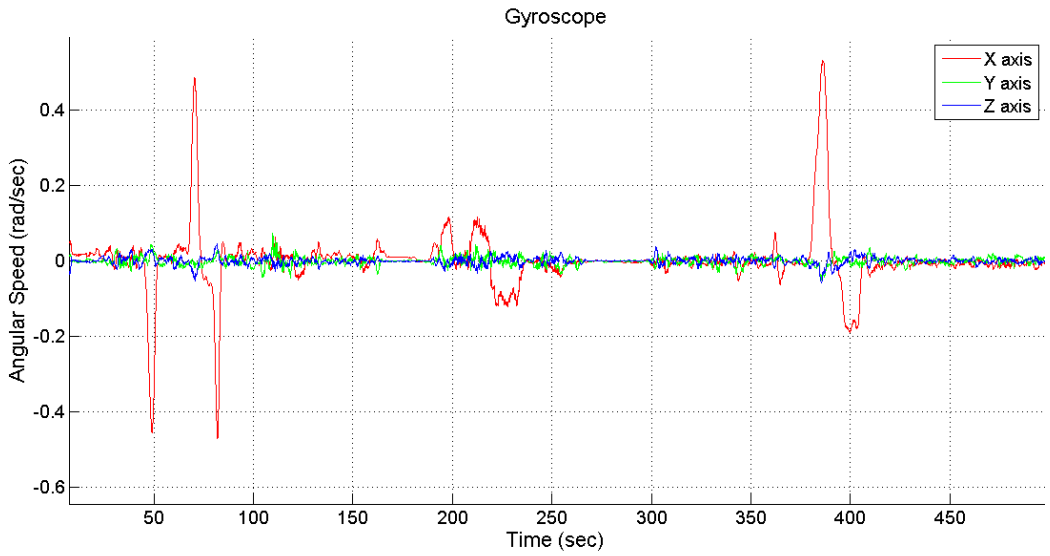


Figure 6.5: Simple Moving Average (SMA) of Gyroscope Values

determining turns involves filtering if the car is drifting, if yes, then whether the drift qualifies as a turn. The thresholds for drifts were determined by visually inspecting the GPS and gyroscope data.

Car qualifies for a drift if the x-axis SMA of gyroscope values consistently remains more than 0.015 rad/sec (0.86°/sec) for longer than 0.5 sec. Whenever SMA drops below the drift threshold of 0.015 rad/sec, the drift is complete. The system detects all such drifts and marks them as potential turns.

A qualifying drift qualifies for a turn if, during the whole during of the drift, the SMA crosses the turn threshold of 0.15 rad/sec (8.6°/sec). If it does, the system qualifies it as a turn and moves on the second step. If the drift doesn't qualify for a turn, it is discarded as the data mostly corresponds to a lane change or a curve on the road.

6.4.2 Combining multiple turns

If there are two turns in quick succession, it is difficult to ascertain angle by GPS coordinates as there is not enough data after the first turn and before the start of second turn. To overcome this limitation, the algorithm combines the two turns if the time between the end of the first turn and start of the second turn is less than 3 seconds. Figure 6.6 shows two turns combined by the algorithm. The black line shows the approximate path of the car and

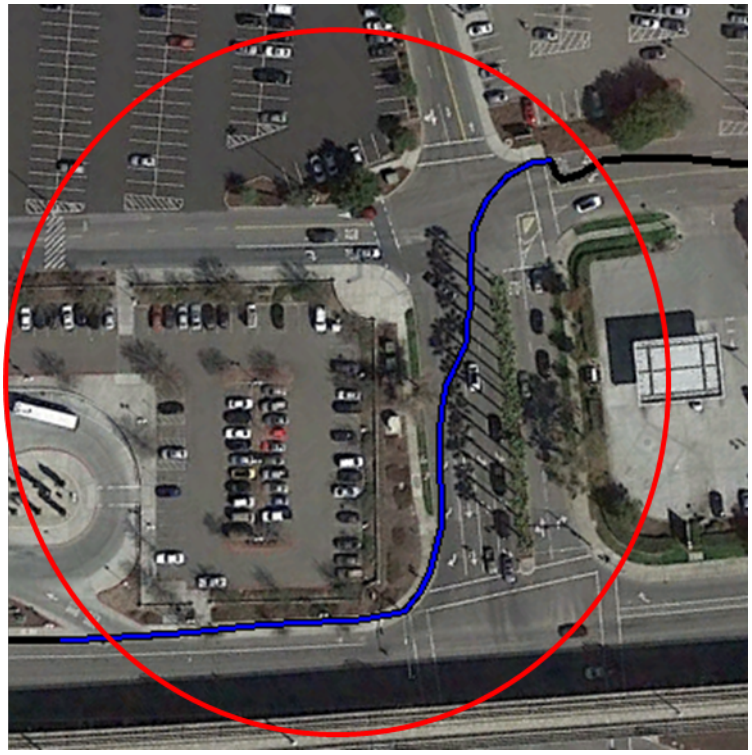


Figure 6.6: Combining two successive turns into one

has been constructed by joining two consecutive GPS locations by a straight line. The blue line represents the GPS locations recorded when the car was turning. It has been constructed by the same method as the black line. Having 3 seconds of data after the turn can ascertain the angle calculated through GPS in most of the scenarios.

6.4.3 Calculating angles

Once the turn has been ascertained, total angle turned by the car is calculated by gyroscope and GPS coordinates.

Integrating gyroscope data along the x-axis over the duration of the turn calculates the total angle covered. There are many methods to perform numerical integration and can be found in [28]. The implementation uses trapezoidal integration.

First, the GPS data needs to be filtered for random offsets as shown in figure 6.7. The easiest way to filter such errors is by setting a maximum speed for the vehicle and eliminating points which are impossible to reach

Figure 6.7: Random offsets

with maximum speed. A maximum speed of 250 km/hr (155 miles/hr) (about 70 meters/sec) eliminates all such invalid points.

Computing angle from GPS is relatively more challenging than from gyroscope data. The angle formed by GPS coordinates are calculated by assuming GPS coordinates on a plane instead of a sphere. Error due to this assumption is small as the processing is done over a small area.

Non-functional technique

An incorrect technique was tried in the first attempt to find the angle of the turn. It aggregated the angles between every two consecutive lines (formed by three consecutive points) over the duration of the turn. Figure 6.8 presents GPS data of a car being driven into a parking lot and then taken out. The red line represents the GPS path of the turn when the car was entering the parking lot and the yellow line represent the turn when the car was exiting the parking lot. As visible on the figure, path formed by GPS coordinates

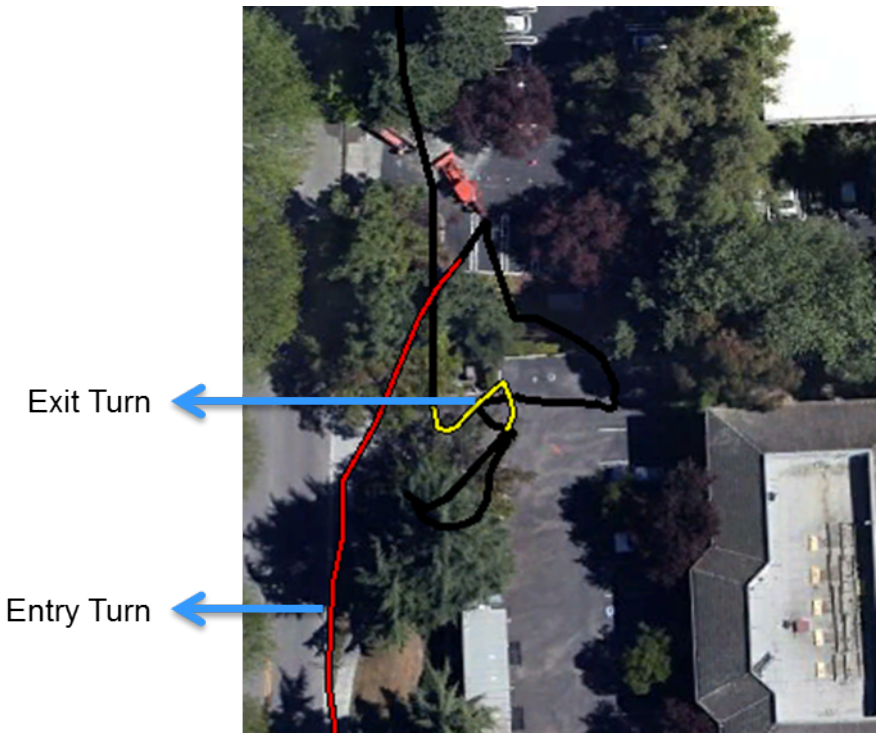


Figure 6.8: Issue with trees and parking lots

changes randomly giving abrupt angles between lines. Sometimes the lines forming the path intersect with each other making angle calculations difficult. It is very difficult to ascertain whether the reflex angle or the angle opposite to reflex angle needs to be taken into account. Selecting incorrect angles leads to incorrect results in all such situations.

Another situation, which is very common at turns is waiting for the signal before the turn (Figure 6.9). In such situations too, the GPS tend to calculate different location, all in nearby places as shown by the circle in 6.9. It makes angle calculation difficult through geometry. In the same figure, the turn depicted by the red line is an example for the turn for which the angle can be calculated accurately with high probability.

Functional technique

Another technique, and the technique used in the final implementation, finds the direction of vehicle travel before and after the turn to measure the angle of the turn. Let **A** and **B** be the unit direction vectors before and after the turn. Angle between them can be found out by

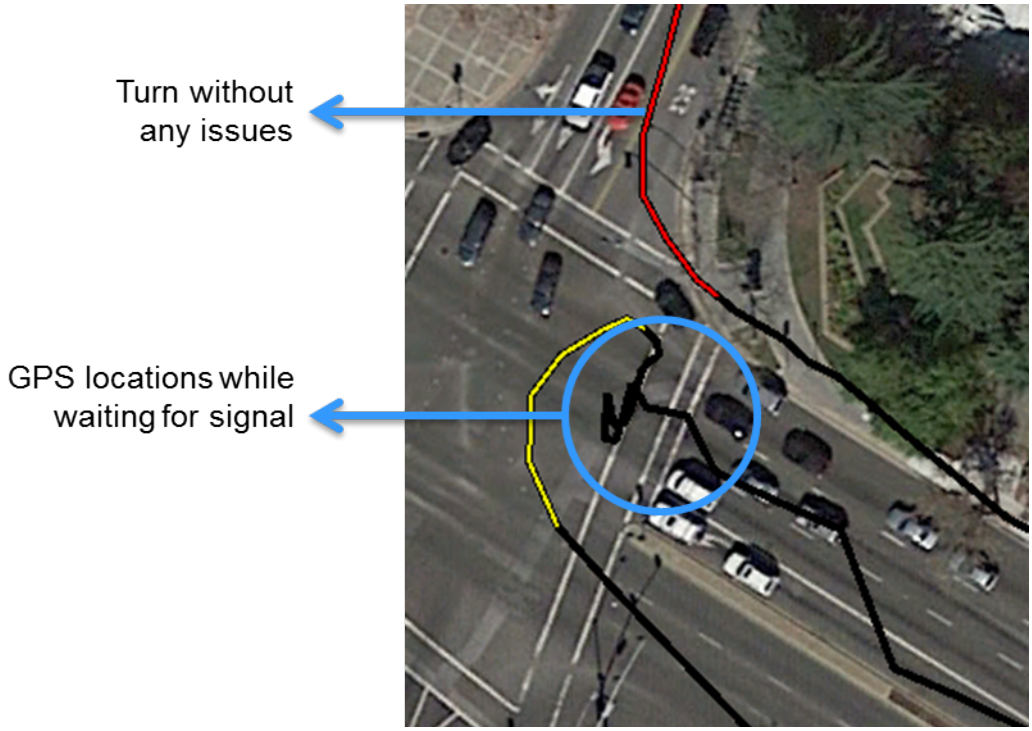


Figure 6.9: Issue at traffic signals

$$\theta = \arcsin(\|\mathbf{A} \times \mathbf{B}\|) \tag{6.1}$$

The result is in the range $-\pi/2$ to $\pi/2$. Integration of gyroscope angle provides the actual angle made by the vehicle and lies outside this range for all the turns making an obtuse angle. Calculating the *sin* equivalent angle of the actual angle in $-\pi/2$ to $\pi/2$ range allows comparison of the angles.

Figure 6.10 shows two turns, a U-turn (turn 1) and a right angle turn (turn 2). By using the above mentioned techniques of detecting turns and GPS angles, angle calculated by gyroscope data for turn 1 comes out to be 179.63°and that by GPS data to be 2.7°. The arcsin equivalent of the gyro angle is 0.37°. It allows for a direct comparison of angles.

Finding direction of travel

To ascertain the direction of travel before the turn, we trace the location coordinates back in time until the distance from the start of the turn more than the threshold of 30 meters. Similarly, for direction after the turn, we trace the location forward in time until we find a location 30 meters away.

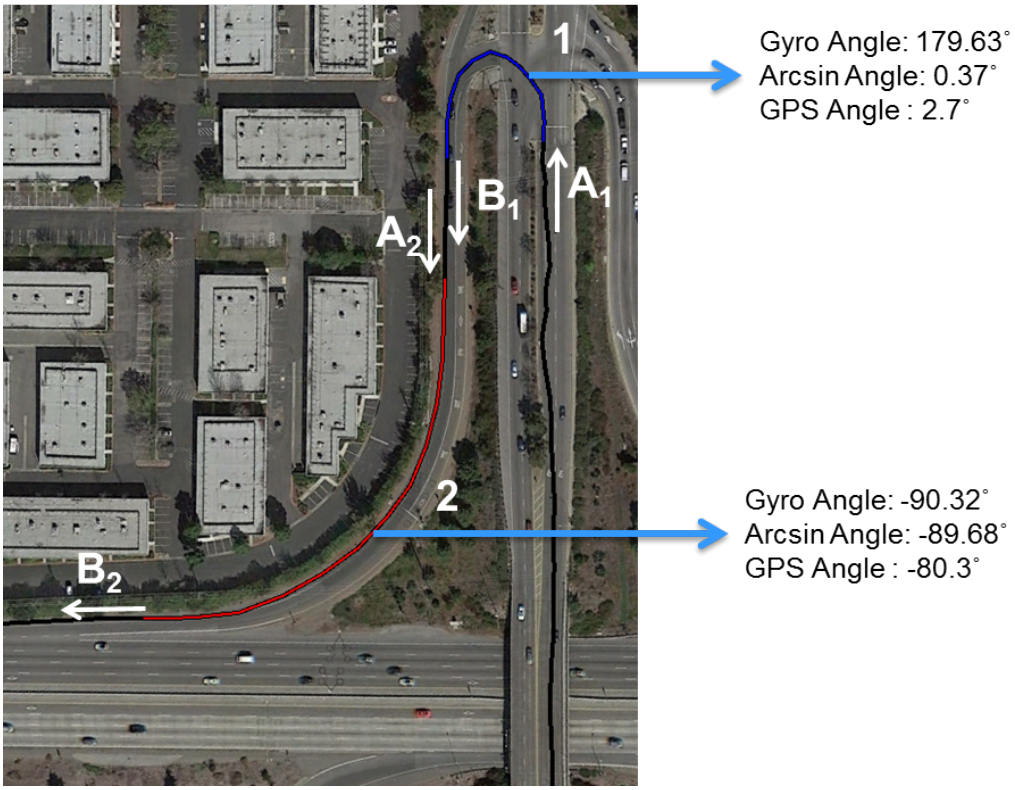


Figure 6.10: Turn angles and arcsin

The direction of travel is the average of the vectors from the start or the end of the turn till the location.

In an open area, GPS has an accuracy of 9-12 meters. 30 meters is 3 times the accuracy to ensure that vehicle is moving in the desired direction.

In figure 6.10, $\mathbf{A_1}$ corresponds to the direction of travel before the turn 1 and $\mathbf{B_1}$ corresponds to the direction after the turn. The length of the vectors in the figure is approximately 30 meters on ground. Same goes for turn 2. In this figure, enough GPS data is available for before and after the turns to get the direction of travel but this may not be the case all the time. The availability, accuracy, and characteristic of the data, changes the confidence in the turn.

6.4.4 Computing the validity score

Validity score determines the confidence in the turn. If the accuracy of the GPS data is high (<15 meters), and enough data is available to determine the direction of travel and turn angle, the validity score is 1.0.

If enough data is not available, for instance, if the system encounters another turn in less than 30 meters, if the car is in a multistory parking garage where GPS is unavailable, or is sporadic with low accuracy because of trees, validity score of the turn goes down.

In the current implementation, validity score goes down when any of the following conditions occur.

- If the distance between any two consecutive turns is less than 30 meters but greater than 9 meters and has high accuracy, validity score is dropped by 0.1. If it has low accuracy (>20 m), we discard the turn with validity of 0. If the distance is less than 9 meters (for example in a parking lot), we discard the turn from validation checking. 9 meters is the highest possible accuracy in Android GPS and any distance less than 9 meters can occur due to inaccuracy.

- If enough data is not available before or after the turn. For instance, in case the recording started or ended when the car was taking the turn, that particular turn is discarded from checking.

- If the total period of turn becomes longer than 10 sec, errors due to integration of gyroscope data become large. The system decreases the validity by 0.1 in this scenario too.

- If the difference between the GPS angle and the gyroscope angle becomes more than 10 degrees, we reduce the validity by 0.3. In other terms, if the available data is not good to ascertain the angle turned, the validity is decreased where the data is bad, besides the decrease when difference in the angle is beyond the threshold. Error can creep in when the turn angle is around 90°. As shown in figure 6.10 for turn 2, if the turn angle is little more than 90°, the system reduces it to less than 90 but close to it. In such cases, the difference in the GPS angle and gyroscope angle will decrease, which might leading to higher validity score than actual.

6.5 Android App

The algorithm described in section 6.4 has been implemented as an Android app, called PPVMT Data Validator.

6.5.1 PPVMT Data Validator

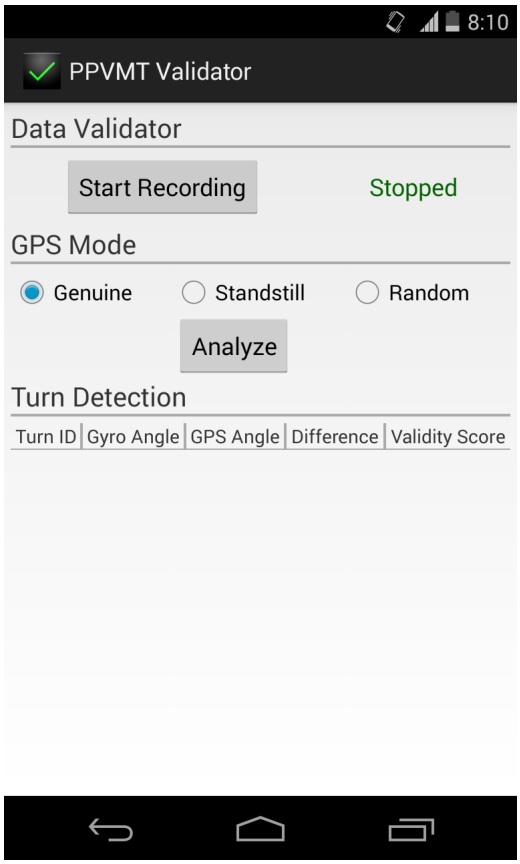


Figure 6.11: Main Screen

Figure 6.11 is the main screen of the app. The app first records the data and then analyzes it for validity. It can analyze the collected data with 3 different GPS modes - Genuine, Standstill, and Random. Genuine mode uses the GPS values as recorded by the app, Standstill feeds the validation algorithm with a static GPS coordinate and Random just chooses any random GPS coordinate.

Start Recording button starts the data recording on the permanent storage as a file. It initiates a background service (Figure 6.12) which spawns a

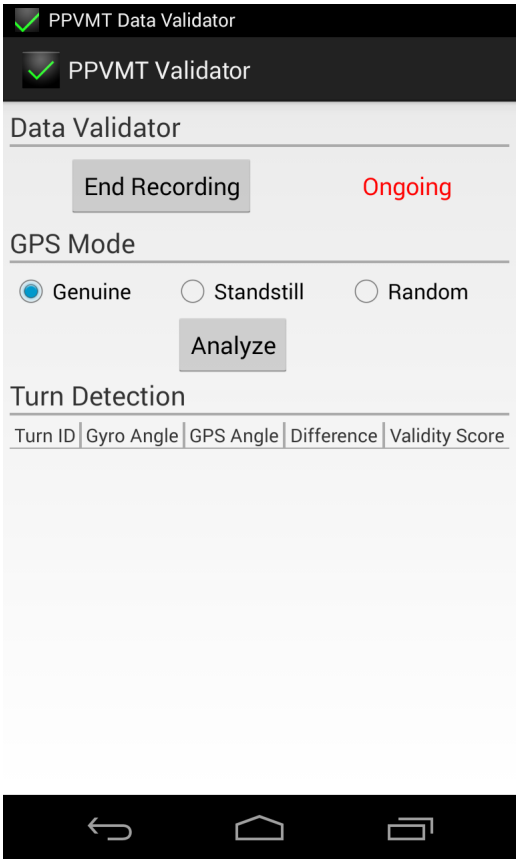
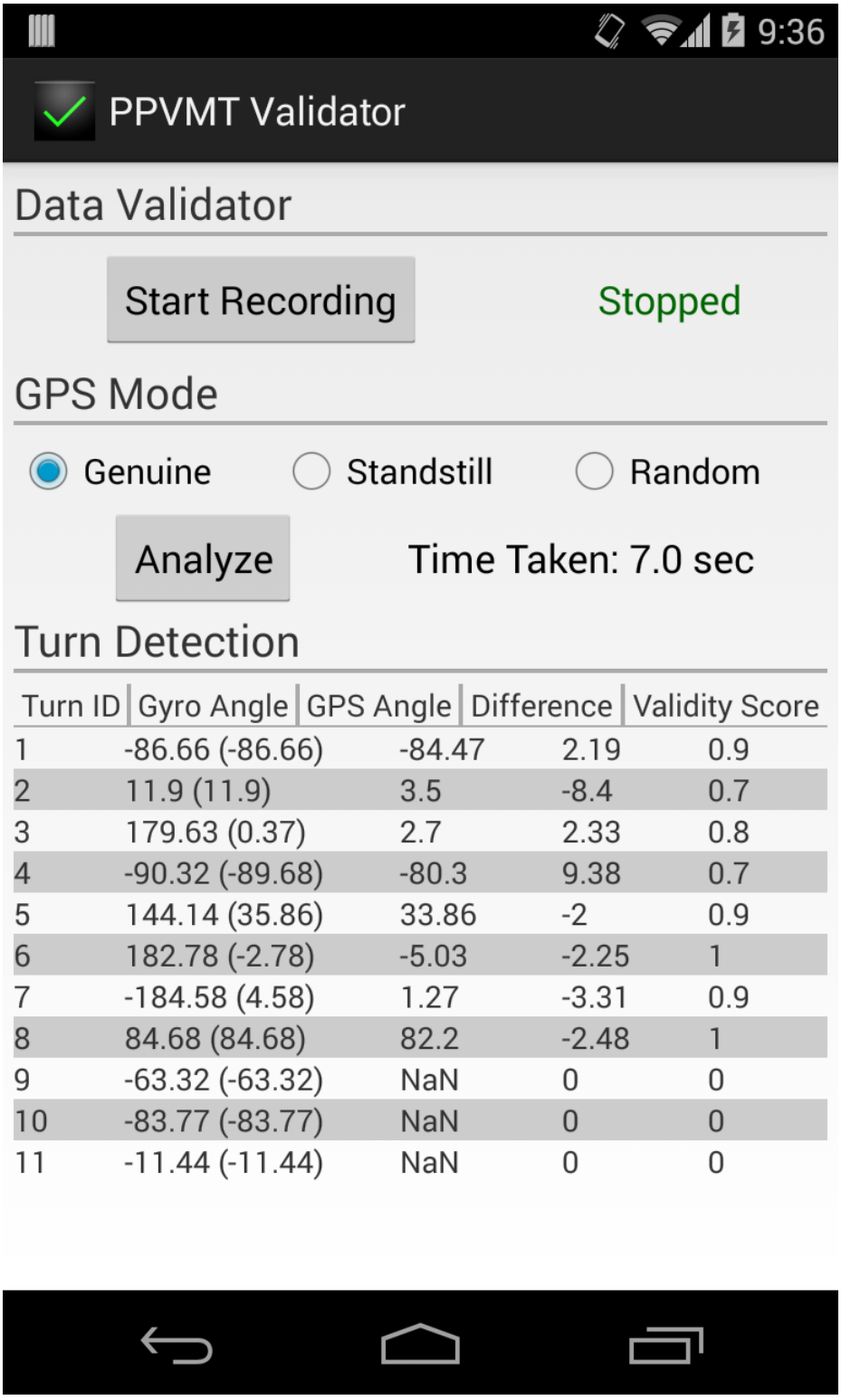


Figure 6.12: Recording data

thread for data collection and writing. Pressing the *End Recording* button in the app stops the recording. *Analyze* button validates the collected data according to the selected GPS mode. *Gyro Angle* is the angle in degrees, calculated by gyroscope data. *GPS Angle* is the angle of the turn calculated using GPS data. *Difference* is the difference in the GPS and gyroscope angle and *Validity Score* is the validity parameter as discussed in section 6.4.

Figure 6.13 shows validation of a 39 minute trip using genuine GPS data. The angle in the bracket next to actual turn angle is the sin equivalent angle in the range of -90°to 90°. This is one of the good data sets where the majority of driving, except the starting turns and the ending turns, occurred on highways. The last three turns are in a multistory parking lot where GPS is unavailable. Overall validity of turns for this data set is really good.

Figure 6.14 shows the algorithm in work with a static GPS location of (37.36754001,-122.01043922). Since the distance covered is always 0.0 meters (<9 meters), all the turns get discarded with every turn have a validity score



Turn ID	Gyro Angle	GPS Angle	Difference	Validity Score
1	-86.66 (-86.66)	-84.47	2.19	0.9
2	11.9 (11.9)	3.5	-8.4	0.7
3	179.63 (0.37)	2.7	2.33	0.8
4	-90.32 (-89.68)	-80.3	9.38	0.7
5	144.14 (35.86)	33.86	-2	0.9
6	182.78 (-2.78)	-5.03	-2.25	1
7	-184.58 (4.58)	1.27	-3.31	0.9
8	84.68 (84.68)	82.2	-2.48	1
9	-63.32 (-63.32)	NaN	0	0
10	-83.77 (-83.77)	NaN	0	0
11	-11.44 (-11.44)	NaN	0	0

Figure 6.13: Data analyzed with genuine GPS coordinates

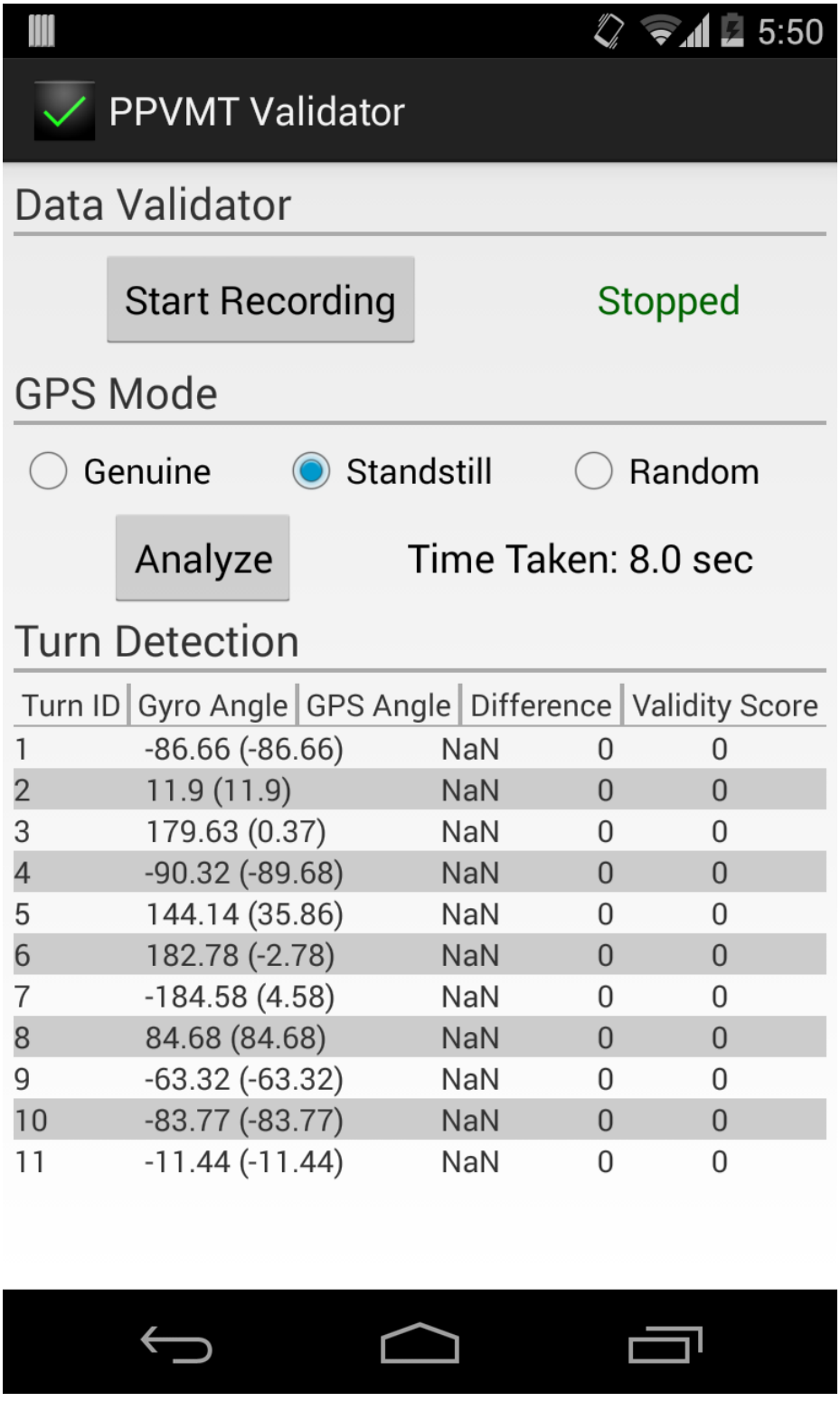


Figure 6.14: Data analyzed with a static GPS coordinate

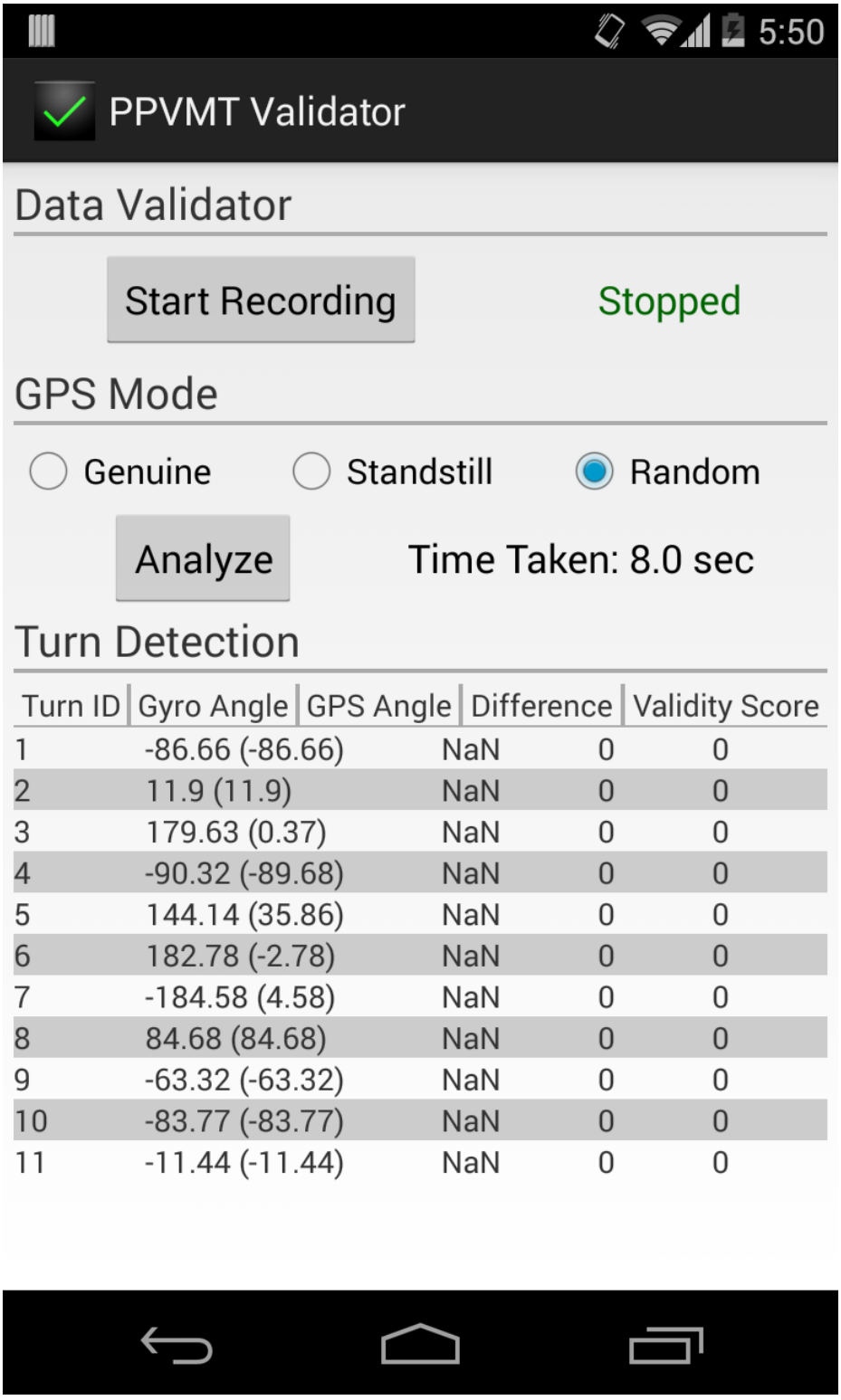


Figure 6.15: Data analyzed with random GPS coordinates

of 0.0.

Figure 6.15 shows the algorithm in work with random generated geographical coordinates. Since the distance between randomly coordinates is really high, all the points get eliminated by the maximum speed filter leading to all turns getting discarded.

Validation of 39 min of driving took 7 seconds on Android Nexus 4 phone. The time taken for processing in standstill and random GPS modes is longer because the algorithm searches extensively for a location with distance more than 30 meters before and after every turn. Only when it is not able to find any, it ignores the turn.

6.6 Performance and Testing

6.6.1 Secret Sharing

United States consists of 3114 counties. Considering each county as a region in itself, we tested the performance of aggregation on Independent Servers and Tax Authority Server by varying the number of cars. The testing machine contained Intel i7-2600 CPU, 12 GB physical memory, Win 7 64 bit and Java 7 update 51.

Figure 6.16 plots the time taken for aggregation on Independent Servers versus the number of cars with 3114 regions. Since all the operations are linearly dependent upon the number of cars, the time required increases linearly.

Figure 6.17 plots the time taken by Tax Authority server for aggregation. Even for 100K cars, the computations took sub-second time.

This test ran on a single thread and didn't utilize multi-core capability. Since the problem is embarrassingly parallel, performing the operations with multiple cores will bring down the processing times significantly. The aggregation can also be performed on highly parallel graphic cards to decrease the processing time drastically.

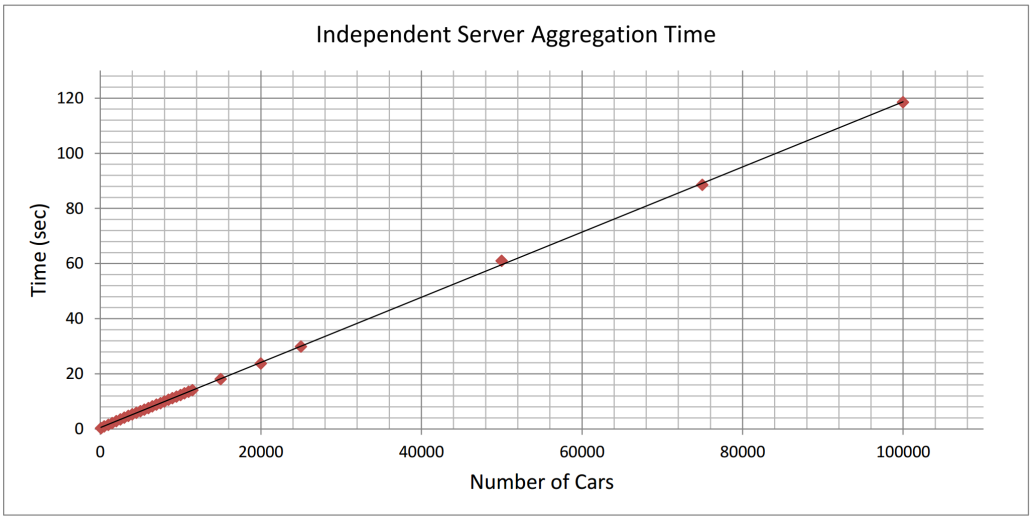


Figure 6.16: Aggregation Performance of Independent Servers

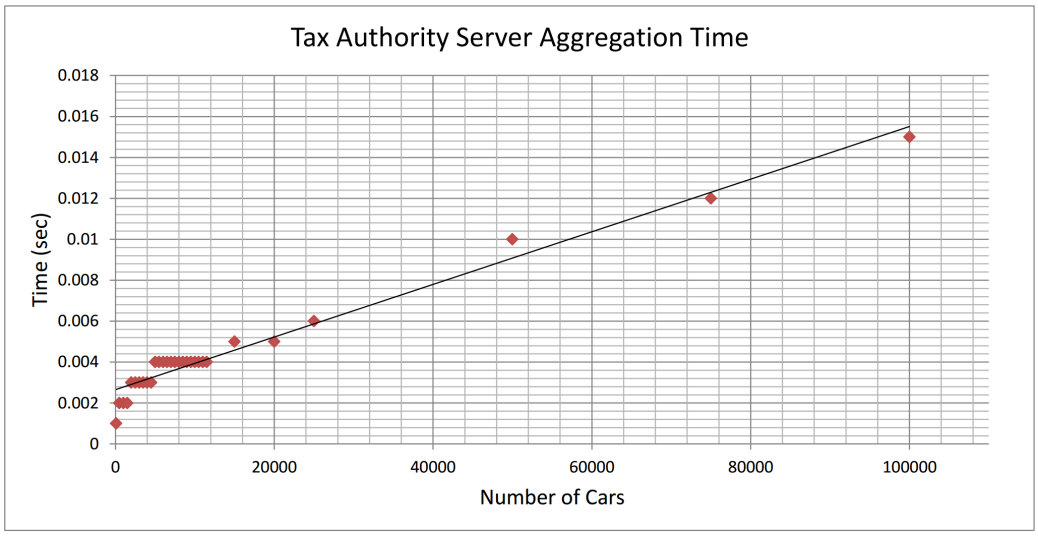


Figure 6.17: Aggregation Performance of Tax Authority

Chapter 7

Related Work

The idea of VMT tax has been there for quite some years now. Governments are actively pursuing various techniques to deploy the tax. Oregon has already passed a Senate Bill for a pilot program, the first in the country. California and Massachusetts are other states where talks on VMT tax are ongoing. In this Chapter we discuss the pilot programs that have been undertaken to test the tax. We also talk about Pay-As-You-Drive (PAYD) insurance, privacy-preserving Smart Meter technologies, and GPS spoofing, all of which relate directly to our work.

7.1 VMT Projects

Charging user based on the region and miles driven has been tested by universities and many government organizations. Some projects aimed to gather data from motorists about their privacy preference and some aimed to develop a proof-of-concept for VMT taxation. Apparently, one of the methods of controlling congestion involves charging users who are in congestion zone. This model of reducing congestion is similar to VMT and has been tested by Puget Sound Regional Council (PSRC) [29].

VMT Pilot Programs by Oregon

Oregon has leaded the states in charging users fuel taxes for maintenance and construction of roads. It was the first state to initiating the fuel tax in 1919. Even today, it leads the nation by passing Senate Bill 810 [8] which allows Oregon Department of Transportation to set up a mileage collection system for 5,000 volunteer motorists beginning in 2015.

In 2001, Oregon formed the Road User Fee Task Force (RUFTF) to exam-

ine alternatives for raising revenue. Oregon's Department of Transportation (ODOT), along with RUFTF conducted two pilot projects—Road User Fee Pilot Program (RUFPP) [30] in 2007 and Road Usage Charge Pilot Program (RUCPP) [31] in 2012. The results of RUCPP were instrumental in passing of the SB 810. These projects were created to charge high efficiency vehicles and not just BVs.

- *Road User Fee Pilot Program (RUFPP)* developed and tested "pay-at-the-pump" approach in 2007 where the users' cars contained on-vehicle device with RF communication and GPS capabilities. At the gas pump, the pumps identified the car by interacting with the device through RF. After identifying the vehicle, the gas pump subtracted the tax amount from the gasoline price. GPS was used to determine the amount of the tax which the user paid later. The test was conducted with 285 subject vehicles but it failed to make it to the legislature because of privacy concerns (mandated GPS), and cost of the required infrastructure. One concern was slow technology evolution if the technology was not subjected to market forces.

- *Road Usage Charge Pilot Program (RUCPP)* With the concern of reducing fuel tax revenues, RUCPP was launched in 2012 [31]. It removed the GPS requirement, allowed private firms and provided many choices for data collection, reporting and payment of the tax. The project targeted high efficiency vehicles with at least 55 miles per gallon. The users were charged a constant tax rate of 1.56 cents per mile. For this pilot, ODOT contracted with Sanef to provide the reporting devices used by the participants and do the data processing. These devices connect to the OBD-II port of the car. The data collected depended on the type of plan user enrolled in [32]. If the user enrolled in the plan where location information was collected, he was exempt from taxes for driving on non-state and private roads. The project concluded in 2013.

  The program provided 5 different plans:

  1. ODOT Flat Rate Plan: This plan didn't depend on the miles driven. The user paid a monthly fee of $45 with no device installed in the car.

2. ODOT Basic Plan: It included an OBD-II port compatible device which provided ODOT with only mileage information.
3. Sanef Basic Plan: This plan is same as ODOT Basic Plan except that the data was sent to Sanef instead of ODOT which allowed users to view and pay their tax directly on the website.
4. Sanef Advanced Plan: This plan used a device with GPS. The location information was sent to Sanef who charged for miles reported only in Oregon.
5. Sanef Smartphone Plan: This plan used motorists' smartphone for location determination and data communication. The users had to install an app developed by Sanef for location determination. It allowed for mix of basic and advanced plans. The user had the power to control when the location is monitored. Miles, for which location was not monitored, were charged assuming they were accumulated in Oregon.

National Evaluation of Mileage-Based Charges

The project was conducted by University of Iowa Public Policy Center [21]. It aimed to assess the technical feasibility and public acceptance of mile based charges with location monitoring. In the 2 year field study, the road charges were examined on the national and multi-jurisdictional scale with approximately 2,650 volunteers from 12 areas throughout the country. The charges were accrued using a On-Board Computer (OBC) consisting of GPS and a connection to OBD-II (On-Board Diagnostics) port of the car for odometer and speedometer data. The computer also stored the region boundaries and rate per mile for jurisdictions. The study accumulated more than 21 million miles with an average of 9000 miles per participant.

The study quantified the GPS ability for determining locations. In the study, 92.5% of the miles was measured successfully. 6.9% of miles required interpolation techniques and 0.6% could not be reliably assigned using straightforward techniques. The study also assessed user acceptance. At the start, more than 60% expressed a neutral or negative view but at the end the system was rated favorable by 70% of the participants.

The paper does not mention any effort to secure the private information

of the users. For privacy protection, it provided users with two options—an invoice containing detailed information about the trips, and another with no data except the total charges. The first option was privacy invasive but allowed audit of the final tax amount. The second option preserved privacy but did not have any audit mechanism. The users had to select between the two and did not have any option of data audit with privacy protection which PPVMT offers.

7.2 Pay-As-You-Drive (PAYD) Insurance

Pay-As-You-Drive (PAYD) insurance is similar to VMT tax as the user is charged based on his driving. In this, the cost of insurance is dependent upon the type, location, time of driving, distance driven, acceleration and speed of the vehicle. The number of factors affecting the cost varies among companies ([33] [34]). The basic idea is to measure the "safeness" of driving. In PriPAYD, authors propose a privacy preserving design for PAYD insurance [22]. It aggregates the data required by insurance company locally on a machine installed in the car and allows only the user to access the fine grained location data. The paper also discusses the privacy invasiveness of GPS data and offers some solutions for tamper resistance. It acknowledges GPS spoofing but does not provide any proof-of-concept.

7.3 Smart Meters

Smart Meters, assessing electricity consumption from power grid, can be privacy invasive too. Fine grained energy consumption data can be used to deduce which appliances are used in a house by a technique known as Non-Intrusive Appliance Load Monitoring (NIALM) [35]. The problem statement in this scenario is similar to VMT tax system - the utility provider needs to receive enough information to make decisions without having access to fine grained consumption data from each household. User privacy can also be preserved by requiring user to do all the computations on the data on his own machine and then send it to the provider with a proof [36]. Danezis et al. [23] computes on encrypted readings, implemented by secret-sharing-

based secure multi-party computation techniques. They also deploy secret-sharing and data aggregation by non-colluding independent authorities.

7.4 GPS Spoofing

Vulnerability of GPS receivers to spoofing has been verified by several tests. In 2013, researchers spoofed GPS signal being received by a $80M yacht and altered it course while the yacht's navigation system showed that the yacht was moving at the correct path [37]. Humphreys et al. implemented a GPS spoofer on a digital signal processor [18]. Numerous countermeasures have been proposed to prevent GPS spoofing [38, 39]. Some of the countermeasures include,

- Checking the power of signals being received on antenna. Spoofed signals have power higher than genuine signals from satellites.
- Checking the time of arrival of signal and ensuring that it lies in the expected range.
- Checking the Doppler shift of the signal frequency.
- Recovering genuine satellite signals by residual analysis. It can be checked with other signals for consistency to detect any spoofing
- Checking the integrity by Receiver Autonomous Integrity Monitoring (RAIM) which uses redundant data from extra satellites. It ignores the measurements which cause large navigation errors.

Chapter 8

Conclusion and Future Work

This thesis described the problem of decreasing motor fuel taxes, leading to a financial deficit for surface transportation infrastructure development and maintenance. Vehicle Miles Traveled (VMT) tax is a reasonable alternative to motor fuel taxes but allocating the tax collected to different governments based on the infrastructure usage require monitoring location of motorists which is privacy invasive.

The thesis proposes an architectural framework, called Privacy-Preserving Vehicle Miles Traveled (PPVMT) tax which protects the location privacy of the motorists. It accumulates total miles traveled in each region and erases any fine grained location. Later, it splits it into values using additive secret sharing and sends it for aggregation on non-colluding servers. The resultant data is then aggregated by the authority responsible for taxation which then calculates the tax owed by each user and the tax to be allocated to each government.

To prevent system tampering, especially, spoofing of GPS and odometer signals, we validate the incoming untrusted data using measurements from inertial motion sensors. We also implemented a proof-of-concept on Android which successfully detected GPS spoofing from gyroscope measurements.

The system has been developed for next generation of cars which will have features similar to smartphone to support cryptographic computations and secure processing. If implemented, the same system can also be used for many different problems related to transportation.

- The system can be integrated with navigation systems similar to Google Maps to allow motorists to choose between paths and minimize the cost of travel.
- Toll collection can be implemented on the same system by considering a section of the road as a region in itself.

- It can be advanced to support temporal pricing which can help transport authorities contain congestion by increasing the prices of regions during high traffic hours. Prices can also be manipulated on the fly to divert traffic and change traffic flows in a city.

- PPVMT can be equipped with techniques to determine safe driving practices, leading to Privacy-Preserving Pay as you Drive (PPPAYD) insurance.

- If the location of a vehicle can be estimated with high precision and accuracy, parking charges can also be collected automatically without invading the location privacy of the user.

VMT tax will remove the reliance on fuel for infrastructure revenue collection and has huge potential to affect the way taxes and other transportation related charges are collected.

References

[1] "Urbana Motor Fuel Tax Rate Change," http://urbanaillinois.us/sites/default/files/attachments/fuel-tax-ordinance.pdf.

[2] "Federal Gasoline and Diesel Tax," http://www.transportation-finance.org/funding_financing/funding/federal_funding/motor_fuel_taxes.aspx.

[3] J. Krumm, "Inference attacks on location tracks," in *Proceedings of the 5th International Conference on Pervasive Computing*, ser. PERVASIVE'07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 127–143.

[4] P. Golle and K. Partridge, "On the anonymity of home/work location pairs," in *Proceedings of the 7th International Conference on Pervasive Computing*, ser. Pervasive '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 390–397.

[5] K. Koscher, A. Czeskis, F. Roesner, S. Patel, T. Kohno, S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, and S. Savage, "Experimental security analysis of a modern automobile," in *Security and Privacy (SP), 2010 IEEE Symposium on*, May 2010, pp. 447–462.

[6] "The Highway Trust Fund and the Treatment of Surface Transportation Programs in the Federal Budget," http://www.cbo.gov/publication/45416.

[7] "New Zealand Road User Charges," http://nzta.govt.nz/vehicle/registration-licensing/ruc/overview.html.

[8] "Oregon Senate Bill 810," https://olis.leg.state.or.us/liz/2013R1/Downloads/MeasureDocument/SB810.

[9] "Apple CarPlay," https://www.apple.com/ios/carplay/.

[10] R. N. Charette, "This car runs on code," http://spectrum.ieee.org/transportation/systems/this-car-runs-on-code.

[11] S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, S. Savage, K. Koscher, A. Czeskis, F. Roesner, and T. Kohno, "Comprehensive experimental analyses of automotive attack surfaces," in *Proceedings of the 20th USENIX Conference on Security*, ser. SEC'11. Berkeley, CA, USA: USENIX Association, 2011, pp. 6–6.

[12] I. Rouf, R. Miller, H. Mustafa, T. Taylor, S. Oh, W. Xu, M. Gruteser, W. Trappe, and I. Seskar, "Security and privacy vulnerabilities of in-car wireless networks: A tire pressure monitoring system case study," in *Proceedings of the 19th USENIX Conference on Security*, ser. USENIX Security'10. Berkeley, CA, USA: USENIX Association, 2010, pp. 21–21.

[13] ISO, *ISO 11898-1:2003 Road vehicles – Controller area network (CAN) – Part 1: Data link layer and physical signalling*, Geneva, Switzerland, 2003.

[14] "OVERSEE," https://www.oversee-project.com/.

[15] H. Bruckner, C. Spindeldreier, H. Blume, E. Schoonderwaldt, and E. Altenmuller, "Evaluation of inertial sensor fusion algorithms in grasping tasks using real input data: Comparison of computational costs and root mean square error," in *Wearable and Implantable Body Sensor Networks (BSN), 2012 Ninth International Conference on*, May 2012, pp. 189–194.

[16] "Retention Periods of Major Cellular Service Providers," https://www.aclu.org/files/pdfs/freespeech/retention_periods_of_major_cellular_service_providers.pdf.

[17] V. N. T. S. Center, "Vulnerability assessment of the transportation infrastructure relying on the global positioning system," Tech. Rep., 2001.

[18] T. E. Humphreys, B. M. Ledvina, M. L. Psiaki, B. W. O'Hanlon, and J. P. M. Kintner, "Assessing the spoofing threat: Development of a portable gps civilian spoofer," in *Proceedings of the 21st International Technical Meeting of the Satellite Division of The Institute of Navigation*, ser. ION GNSS 2008, 2008, pp. 2314–2325.

[19] "Odometer Fraud," http://www.carfax.com/car_buying/odometer.cfx.

[20] "Smelecom," http://smelecomuk.com/.

[21] P. F. Hanley and J. G. Kuhl, "National evaluation of mileage-based charges for drivers," in *Transportation Research Record: Journal of the Transportation Research Board*, ser. Volume 2221 / 2011 Revenue, Finance, and Economics. Transportation Research Board of the National Academies, 2011, pp. 10–18.

[22] C. Troncoso, G. Danezis, E. Kosta, J. Balasch, and B. Preneel, "Pripayd: Privacy-friendly pay-as-you-drive insurance," *Dependable and Secure Computing, IEEE Transactions on*, vol. 8, no. 5, pp. 742–755, Sept 2011.

[23] G. Danezis, C. Fournet, M. Kohlweiss, and S. Zanella-Béguelin, "Smart meter aggregation via secret-sharing," in *Proceedings of the First ACM Workshop on Smart Energy Grid Security*, ser. SEGS '13. New York, NY, USA: ACM, 2013, pp. 75–80.

[24] R. Anderson, *Security Engineering*. Wiley, 2008.

[25] "AUTOSAR," http://www.autosar.org/.

[26] "GENIVI Alliance," http://www.genivi.org/.

[27] "Open Automotive Alliance," http://www.openautoalliance.net/.

[28] "Numerical Integration Techniques," http://mathworld.wolfram.com/NumericalIntegration.html.

[29] "Traffic Choices Study, PSRC," http://www.psrc.org/transportation/traffic.

[30] "Road User Fee Pilot Program (RUFPP)," http://www.oregon.gov/ODOT/HWY/RUFPP/docs/rufpp_finalreport.pdf.

[31] "Road Usage Charge Pilot Program (RUCPP)," http://www.oregon.gov/ODOT/HWY/RUFPP/Pages/index.aspx.

[32] "Road Usage Charge Pilot Program (RUCPP) Final Report," http://www.oregon.gov/ODOT/HWY/RUFPP/docs/RUCPP%20Final%20Report%20-%20May%202014.pdf.

[33] "Progressive Snapshot," http://www.progressive.com/auto/snapshot/.

[34] "The Hartford TrueLane," http://www.thehartford.com/auto-insurance/truelane-savings.

[35] G. Hart, "Nonintrusive appliance load monitoring," in *Proc. of the IEEE*, vol. 80, no. 12, Dec. 1992, pp. 1871 –1891.

[36] A. Rial and G. Danezis, "Privacy-preserving smart metering," in *Proceedings of the 10th Annual ACM Workshop on Privacy in the Electronic Society*, ser. WPES '11. New York, NY, USA: ACM, 2011, pp. 49–60.

[37] "Spoofing Yacht," http://www.utexas.edu/news/2013/07/29/ut-austin-researchers-successfully-spoof-an-80-million-yacht-at-sea/.

[38] H. Wen, P. Y.-R. Huang, J. Dyer, A. Archinal, and J. Fagan, "Countermeasures for gps signal spoofing," in *Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation*, ser. ION GNSS 2005, 2005, pp. 1285–1290.

[39] J. S. Warner and R. G. Johnston, "Gps spoofing countermeasures," in *Homeland Security Journal*, 2003.