Fueoogle: A Participatory Sensing Fuel-Efficient Maps Application

Nam Pham, Raghu Ganti, Saurabh Nangia, Thadpong Pongthawornkamol, Shameem Ahmed, Tarek Abdelzaher, Jin Heo, Maifi Khan, and Hossein Ahmadi

Department of Computer Science, University of Illinois, Urbana-Champaign

nampham2, rganti2, nangia1, tpongth2, ahmed9, zaher, jinheo, mmkhan2, hahmadi2@illinois.edu

Abstract

This paper develops a participatory sensing service, called *Fueoogle*, that maps vehicular fuel consumption on city streets, allowing drivers to find the most fuel-efficient routes for their vehicles between arbitrary end-points. The service exploits measurements of vehicular sensors, available via the OBD-II interface that gives access to most gauges and engine instrumentation. The OBD-II sensors are standardized in all vehicles produced in the US since 1996, constituting some of the largest "sensor deployments" to date. Using fuel-related measurements contributed by participating vehicles, we develop a route planner that maps normalized fuel-efficiency of city streets, enabling vehicles to compute minimum fuel routes from one point to another. Street congestion, elevation variability, average speed, and average distance between stops (e.g., stop signs) lead to changes in the amount of fuel consumed making fuel-efficient routes potentially different from shortest or fastest routes, and a function of vehicle type. Our experimental study answers two questions related to the viability of the new service. First, how much fuel can it save? Second, can it survive conditions of sparse deployment? The main challenge under such conditions is to generalize from relatively sparse measurements on a subset of streets to estimates of measurements of an entire city. Through extensive experimental data collection and evaluation, conducted over the duration of a month across several different cars and drivers, we show that significant savings can be achieved by choosing the right route. We also provide extensive results pertaining to the accuracy of models that are used for prediction of fuel consumption values.

1 Introduction

An emerging category of sensor network applications [8, 1, 19, 10, 18] rely on data collection by individuals and sharing of this data within a community for common purposes such as mapping of physical phenomena or computing community wide statistics. In this paper, we develop a novel participatory sensing application, called *Fueoogle*, that computes fuel efficient routes from one point to another. Fueoogle relies on data collected by individuals from their vehicles as well as on the mathematical models that we develop in this paper to compute fuel efficient routes.

Vehicles that have been sold in the United States after 1996 are mandatorily equipped with a sensing subsystem called the *On-Board Diagnostic* (OBD-II) system. The OBD-II is a diagnostic system that monitors the health of the automobile using sensors that measure approximately 100 different engine parameters. Examples of monitored measurements include fuel consumption, engine RPM, coolant temperature, and vehicle speed. A comprehensive list of measured parameters can be obtained from standard specifications as well as manufacturers of OBD-II scanners [3]. Several commercial OBD-II scanner tools are available [3, 4, 2, 5], that can read and record these sensor values.

Fueoogle utilizes a vehicle's OBD-II system and a typical scanner tool in conjunction with a participatory sensing framework to develop a novel application that enables the reduction of fuel consumption of vehicles in everyday use (by computing fuel efficient routes). Compared to traditional mapping tools, such as Google maps [15] and MapQuest [22], which provide either the fastest or the shortest route between two points, Fueoogle collects the necessary information to compute and answer queries on the *most fuel-efficient route*. The most fuel-efficient route between two points may be different from the shortest and fastest routes. For example, a fastest route that uses a freeway may consume more fuel than the most fuel-efficient route because fuel consumption increases non-linearly with speed or because it is longer. Similarly, the shortest route that traverses busy city streets may be suboptimal because of downtown traffic. The optimal route might therefore be neither shortest nor fastest. Indeed, we will show, in this paper, examples where the most fuel-efficient route is different from both the shortest and the fastest routes.

The motivation for Fueoogle does not need elaboration. Fueoogle users might be driven by benefits such as saving on fuel or reducing CO_2 emissions and the carbon footprint. With the increase in the use of bluetooth devices (e.g., cellphones) and in-vehicle Wi-Fi, Fueoogle can be easily supported by inexpensive OBD-II-to-bluetooth or OBD-II-to-

WiFi adaptors that can upload OBD-II measurements opportunistically, for example, to applications running on the driver's cell phone. It can also be supported by scanning tools that read and store OBD-II measurements on storage media such as SD cards. At the time of writing, OBD-II Bluetooth adaptors, such as the ELM327 Bluetooth OBD-II Wireless Transceiver Dongle, are available for approximately $50, together with software that interfaces them to phones and handhelds. Individuals who own OBD-II adaptors or scanning tools may record sensor measurements from their daily commutes. These recorded sensor values are then shared within a community, in a privacy-preserving fashion, using a participatory sensing framework, called PoolView [13]. Fueoogle does not require all city streets to be driven by all types of vehicles in order to estimate the fuel efficiency of different vehicle types on different streets. Instead, Fueoogle utilizes models, we develop in this paper, to estimate fuel consumption for different streets and car types, for which no direct OBD-II measurements are present, using previously collected data on other streets and car models.

Fueoogle supports two types of users; members and non-members. Members are those who contribute data to the Fueoogle repository from OBD-II sensors as described above. They have Fueoogle accounts and can benefit from more accurate estimates of route fuel-efficiency, customized to the performance of their individual vehicles. Non-members can use Fueoogle to query for fuel-efficient routes as well. Since Fueoogle does not have measurements from their specific vehicles, it answers queries based on the average estimated performance for their vehicle's make and model. In addition to being a look-up service such as Google Maps, the authors envision Fueoogle to be integrated as an option in future "green" GPS services that would give directions based of the most fuel-efficient (as opposed to the fastest or shortest) route.

In summary, the main contributions of this paper are two-fold. First, we develop a fuel-saving service and analyze the amount of fuel savings that are achieved using Fueoogle. An experimental study is performed over the course of a month using seven different cars with different drivers in order to estimate fuel savings. The second contribution, and the main challenge addressed in this paper, is whether we can use a sparse deployment to estimate the fuel consumption on streets and car types for which OBD-II measurements are not yet available. We develop several mathematical models, using the datasets obtained over the course of our experimental study, to correlate fuel efficiency with observable parameters such as street speed limits, presence and number of traffic lights, congestion information, and the type of car for which the route is computed (e.g. SUV, small sedan).

The rest of this paper is divided into six sections. Section 2 presents a feasibility study that investigates the amount of fuel savings that can be achieved by using Fueoogle and by following the fuel-efficient routes. The details of Fueoogle system are described in Section 3. Models for estimating fuel consumption on streets lacking such measurements are presented in Section 4. Evaluation results are presented in Section 5. Related work is presented in Section 6. Finally, we conclude with directions for future work in Section 7.

2 A Feasibility Study

In this Section, we present a feasibility study that provides the reader with an estimate of fuel savings that can be achieved by driving on the most fuel efficient routes.

We compute fuel consumption between landmarks (in the city where the authors reside[1]) and compare these values across multiple routes between the same pairs of landmarks. The landmarks chosen were frequently visited destinations such as the work place of the authors, a major shopping center, and a football stadium. Figure 1 shows the routes used in the experiments. Each experiment was performed independently from the other experiments using multiple cars and different users. For the purposes of Experiment 1 and Experiment 3, we used data collected from a Pontiac Grand AM, 1997. For Experiment 2, the car used to collect data was a Honda Civic, 2002. The shortest and fastest paths are obtained using commonly available mapping services such as Google maps [15] and MapQuest [22] [2]. We plot the fuel consumption for the shortest path, the fastest path, and the path that consumes the least fuel for these three experiments in Figure 2.

We observe, from Figure 2, that in the first experiment, the fastest path is also the most fuel efficient path. Whereas, in the second experiment, the shortest path consumes the least amount of fuel. In the third experiment, the most fuel-efficient route is different from both the shortest and the fastest routes (which happen to be the same). We conclude from the above observations that simply choosing the shortest or the fastest path will not necessarily result in the most fuel-efficient path.

The most conservative estimate of fuel savings obtained from the experiments shown in Figure 2 is about 10% (and the average is 15% across all the three experiments). At the current national average gas price (which is about $2), this would be equivalent to a savings of at least 20 cents per gallon at the pump, which is not bad for "cash back".

To estimate the amount of savings that can be achieved on a global scale, we provide back of the envelope calculations based on data from the Environmental Protection Agency (EPA) [11]. An estimated 200 million light vehicles (passenger cars and light trucks) are on the road in the US. Each of them is driven, on an average, 12000 miles in a year. The average mile-per-gallon (mpg) rating for light vehicles is 20.3 mpg. Even if 5% of these vehicles adopted Fueoogle and the 10% fuel savings were achieved on only a quarter of the routes traveled by each of these vehicles, the amount of overall fuel savings is nearly 148 million gallons of fuel $((12000*0.25)/20.3*(0.05*200M)*0.1)$. This translates into about one third of a billion dollars in savings at the pump (based on the current national average pump prices for a gallon of gasoline). The authors consider the above prospective savings acceptable. The rest of the paper presents details of the Fueoogle service.

[1] City name is removed for anonymity

[2] Google maps provides only the shortest path, MapQuest provides both fastest and shortest paths, hence we use MapQuest to get route information

(a) Figure showing the driving routes used in Experiment 1 (from point A to point B) and Experiment 2 (from point C to point D). Routes with dashed and dotted lines are shortest-path routes while routes with dash lines are fastest routes provided from MapQuest system. The most fuel efficient route is marked by solid lines. Experiment 1 was conducted using a Honda Civic, 2002 and Experiment 2 using a Pontiac Grand AM, 1997.

(b) Figure showing the example of the most fuel-efficient route that is different from the fastest and shortest-path route. The route with solid lines is the fastest and shortest-path route from point A to point B provided by MapQuest system. The route with dash lines is the most fuel-efficient route. The car used for collecting data was a Pontiac Grand AM, 1997.

Figure 1. Maps showing the experiments performed for the feasibility study

3 The Fueoogle System

The service provided by Fueoogle is similar to a regular map application, such as Google maps [15] or MapQuest [22]. Google maps and MapQuest provide the shortest or fastest routes between two points, whereas Fueoogle computes the most fuel-efficient route. A snapshot of the Web-based Fueoogle's user interface is shown in Figure 3 along with the most fuel efficient route between two points for a user with a Pontiac Grand AM, 1997. In the following subsections, we will discuss the Fueoogle concept, then present the participatory sensing framework that we utilize for data collection and data sharing and the specifics of the hardware used for the purpose of data collection.

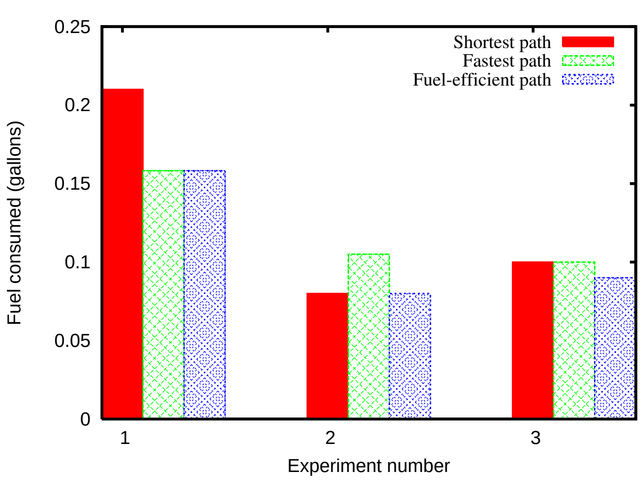


Figure 2. Figure showing fuel consumption for multiple routes between multiple selected landmarks for different cars and drivers

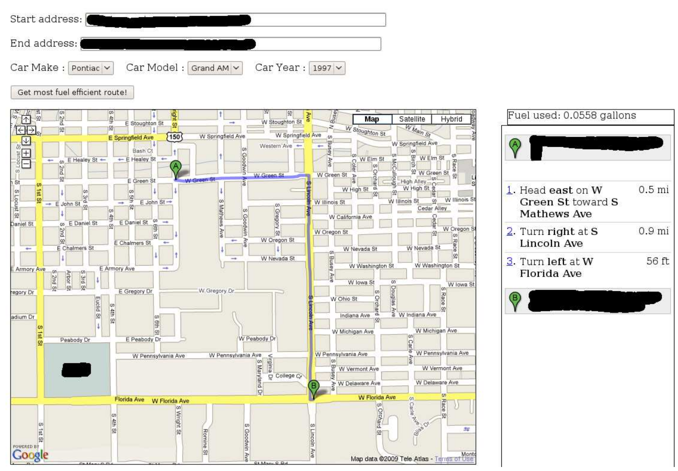


Figure 3. Figure showing the user interface of Fueoogle with the most fuel efficient route between two points on the map for a Pontiac Grand AM, 1997 car model

3.1 The Fueoogle Concept

Individuals who want to compute the most fuel-efficient route between two points enter the source and destination address via the interface provided by Fueoogle. Members of

Fueoogle (i.e., those individuals who contributed participatory data) can register their vehicles that were used for data collection. Hence, Fueoogle can compute the route specifically for the registered vehicle. Other users may enter their vehicle's make, model, and year of manufacture, as well as average mpg and tire diameter, if known (for better accuracy). Alternatively, those last two parameters can be looked up from the model data. Since different vehicles have different fuel consumption characteristics, these car details are used to compute the most fuel-efficient route for the given vehicle brand. The advantage for the users who contribute data is that the system provides better estimates of the most fuel-efficient routes to these individuals, thus allowing them to have higher savings. This is because the prediction models (Section 4) will be more accurate for those individuals who contribute data on their specific cars.

It is impractical to assume that Fueoogle members will measure all city streets and cover all vehicle types. Instead measurements of Fueoogle members are used to calibrate generalized fuel-efficiency *prediction models*. These models, discussed in Section 4, show that the fuel consumption on an arbitrary street can be predicted accurately from set of *static* street parameters (e.g., the speed limit, the number of traffic lights, and the number of stop signs) and a set of *dynamic* street parameters (such as the average speed on the street or the average congestion level), plus of course the vehicle type (specifically, its mpg rating and some known parameters such as wheel diameter and weight). It is the mathematical model describing the relation between these general parameters and fuel-efficiency that gets estimated from participant data. Hence, the larger and more diverse is the set of participants, the better the generalized model.

For most streets, static street parameters can be readily obtained from traffic databases. For example, the number of traffic lights, the number of stop signs, and the speed limits of streets can be obtained from the red light database [16]. Dynamically changing parameters such as the congestion levels or average speed are more tricky to obtain. In larger cities, real-time traffic monitoring services can supply these parameters [25]. Many GPS device vendors, such as TomTom, also collect and provide congestion information. Finally, participatory sensing applications, such as Traffic Analyzer [13] and CarTel [19], have been described in prior literature that have the potential to provide congestion and speed data. In this paper, we use historic per-street-block traffic speed averages computed by the Traffic Analyzer participatory sensing service. Fueoogle utilizes this service for (historic) average congestion level information. The Traffic Analyzer archives these averages for different city blocks as a function of the time of day and day of the week, based on GPS data collected from individuals with GPS devices (that are much more common than the OBD-II scanners). Hence, while Fueoogle is not yet responsive to real-time conditions, such as accidents on the road, it can still provide information on which of multiple routes is the most fuel-efficient *on average* at a given time and on a given day.

3.2 A Participatory Sensing Framework

We utilize a participatory sensing framework, called PoolView [13], to implement Fueoogle. Briefly, PoolView is a set of infrastructure tools and protocols that enable individuals to set up new participatory sensing services and collect data for them. PoolView consists of four layers; namely *sensing*, *storage*, *privacy*, and *aggregation*. The *sensing* layer encompasses participants' sensors. It includes drivers that allow them to upload data to the user's private archive. The private *storage* layer maintains the archive of sensory data collected by the user. The *privacy firewall* layer implements various privacy policies to sanitize data by masking or perturbing appropriate fields prior to sharing with external participatory sensing services. The *aggregation* layer implements such services that aggregate sanitized data and compute service-specific statistics. The communication between these layers is achieved using an extension to the HTTP protocol.

We implemented Fueoogle as a participatory sensing service (i.e., an aggregation server) in PoolView. An individual who wants to share their OBD-II sensor data can thus download the client side software of PoolView, and use it to upload their data to the Fueoogle aggregation server. The aggregation server uses these data to calibrate models that relate street and vehicle parameters to fuel-efficiency and offers the Fueoogle query interface for fuel-efficient routes.

Individuals who wish to contribute OBD-II data to Fueoogle can install, in their vehicle, any commercial OBD-II scanner along with a GPS unit. In our deployments, we use one such off-the-shelf device for data collection purposes. Our hardware setup consists of an OBD-II scanner connected to a GPS unit, as shown in Figure 4. We use DashDyno's OBD-II scanner [3] for collecting sensor data from a car and a Garmin eTrex Legend GPS [14] to get location data. DashDyno has a GPS port allowing the Garmin to be plugged in. The DashDyno records trip data (including Garmin's GPS location) on an SD card that the user later uploads to the Fueoogle server.

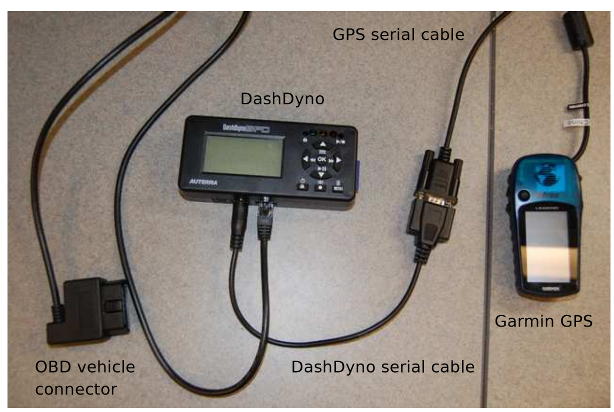


Figure 4. Hardware setup used for data collection

A total of 19 parameters are obtained from the car and the GPS, the most important of them being instantaneous vehicle speed, total fuel consumption, rate of fuel consumption, latitude, longitude, and time. The following section elaborates on the implementation of the Fueoogle server.

3.3 Implementing the Fueoogle Server

The aggregation server provides a fuel-efficient route computation service based on models generated from the data collected. The aggregation server maintains a city's map as a directed graph, which we call the *street graph*, with street intersections as nodes and streets as arcs. The graph is directed because some streets are one-way and because fuel efficiency may depend on the direction of motion. For example, on uneven ground, one direction can be up-hill, while the other down-hill, making fuel efficiency different in each direction[3]. Each arc is assigned a set of parameters such as the speed limit and whether the arc contains traffic lights or stop signs. We call these the *street parameters*. A *fuel consumption model* describes how to map these parameters to consumed fuel.

The server allows registering user vehicles by members. From the data uploaded for each registered vehicle (namely, streets traversed and fuel consumed), a vehicle-specific model is found by regression that maps street parameters to that vehicle's fuel consumption. The purpose of computing such a model is that it can then be used to predict the vehicle's fuel consumption on streets that this vehicle has not traveled from street parameters.

In addition to computing vehicle-specific fuel consumption models for individual registered vehicles, the data on registered vehicles are aggregated into progressively larger pools, that are classified, respectively, by (i) make, model, and year, (ii) make and model, and (iii) make and type (e.g., "compact", "economy", "midsize", "full", etc), and (iv) type. A model is then found for data in each pool. These more general fuel consumption models are used to predict fuel-efficient routes for vehicles that are not registered with Fueoogle. Given the make, model and year of such a vehicle, Fueoogle finds the most specific fuel consumption model that matches that information, then uses it for prediction. For example, to compute a fuel-efficient route for a Ford Taurus, 2001, that is not registered, the server first checks to see if it has a fuel consumption model for a Ford Taurus, 2001, obtained from registered vehicles; if not, then a model for Ford Taurus in general; if not, then a model for full-size Ford vehicles. If not, then a model for full-size vehicles in general. As more vehicles register, models of narrower categories get populated, but the generalizations are useful for early phases of deployment.

Importantly, by parameterizing the vehicles themselves (e.g., by mpg and tire diameter), Fueoogle is able to come up with accurate models that estimate the fuel consumption of one vehicle given data collected by another vehicle. These models are especially good for accounting for finer differences between vehicles in one category (e.g., full-size vehicles) but can also be applied across categories, as will be shown in the evaluation.

When a query is posed for a fuel-efficient route from one point to another in the city, the source and destination addresses of the query are translated into nearest nodes in the street graph. The most fuel-efficient route is the weighted shortest path between these nodes of the street graph, where the weights represent total fuel consumption on each arc, computed using the arc's street parameters and the model used for the vehicle in question. This shortest path is computed using the weighted Dijkstra's algorithm. It is straightforward to extend the above algorithm when the source or destination addresses, or both, are not nodes (i.e., not street intersections). The fuel consumption for segments that represent partial arcs is approximated by multiplying the fuel consumption for the arc by the ratio of the length of the segment to total arc length.

Members can upload more data on their vehicles to the server at any time. The latitude and longitude (location) information shared in conjunction with fuel consumption is used to infer the street it corresponds to based on the *shapefiles* from the TIGER database [26]. The TIGER shapefiles are spatial extracts from the US government's census bureau database which contain feature information regarding the latitude and longitude of various streets/roads in a city. Apart from these features, the database also contains railroad, river, and points of interest information. The fuel consumption and the street information are used to compute the model as detailed below.

4 Prediction Model

One of the main contributions of this paper is to develop a mathematical model that predicts the fuel consumption on streets for which OBD-II measurements are unavailable. Although a large number of people own cars, not many of them have OBD-II scanner tools. The lack of widespread availability of these scanner tools implies that the data being contributed by the users of our participatory sensing application may be rather sparse. Hence, a primary research question is whether one can derive good models for predicting fuel consumption under conditions of sparse deployment. In other words, can we use data collected by a smaller population to build a model that is capable of predicting the fuel consumption characteristics of those streets for which OBD-II measurements are not available? In addition, the different cars have different fuel economy factors that have great effect on the fuel efficiency. Thus the model should be able to accurately predict the fuel consumption for different car models.

There are several factors that affect the fuel consumption on streets. We classify these parameters into four categories, that are (i) *static street parameters*, (ii) *dynamic street parameters*, (iii) *car specific parameters*, and (iv) *personal parameters*. Static street parameters model the street characteristics and do not change (or change very infrequently) over a period of time. For example, the speed limits of streets change very infrequently and the number of traffic lights on the street remain more or less constant. The dynamic street parameters are characteristics that change with time. For example, the congestion levels on a street or the average speed on a street. The static and dynamic street parameters together determine the fuel efficiency of a particular street. Other variations in the fuel consumption can occur due to the type

[3]While the authors appreciate the importance of accounting for street incline as a model parameter, this study does not investigate the effect of incline due to the flat nature of the terrain in the locale where the study is performed.

of car being driven and the nature of the person's driving. For example, a big car may consume more fuel than a small sedan. Similarly, a person who is more erratic (higher acceleration or hard braking) is likely to consume more fuel than a more "careful" driver. These parameters account for the variation in fuel consumption due to the car type and the driver behavior.

Before we explain the details of the model, we provide a brief description of the data collection for the purpose of developing models.

4.1 Data collection

Our model is derived using data collected from six users (with different cars) over the course of a month. A wide range of cars were used in our experiments and a total of about 90 miles were driven by the users. The details of the car make, model, year, and the number of miles of data collected for each car are summarized in Table 1.

Car make	Car model	Car year	Miles driven
Pontiac	Grand Prix	1997	24.5
Honda	Civic	2002	10.55
Chevrolet	Prizm	1998	15.5
Ford	Taurus	2001	9.46
Mazda	626	2001	8.89
Hyundai	Santa Fe	2008	21.4

Table 1. Table summarizing the cars used and the amount of data collected

In our experiments, each user was given a DashDyno and GPS system described in Section 3.2 and was asked to drive around the city in which the authors reside. There were two sets of experiments performed by us, one is a controlled set of experiments, which enabled us to collect sufficient data for a variety of streets. In these set of experiments, each user was asked to drive around a specific set of major streets in the city. Each street had various characteristics, such as the speed limit, the congestion levels, and the number of traffic lights. These controlled experiments captured the variables affecting fuel consumption. The controlled experiments allow us to decide on the best model *structure*, as opposed to estimating parameter values. They are done only once for purposes of understanding what parameters to monitor, and are not part of the participatory service itself. That service will simply use the model structure we arrive at in this paper and use participant data to estimate model parameter values.

The data from these controlled experiments were used to build models. Parameters of these models were estimated. The second set of experiments evaluated the efficacy of the models in an uncontrolled setting. The users drove randomly over several streets and collected (ground truth) fuel consumption information for these streets. The fuel consumption data for these streets were compared against Fueoogle predictions and hence used to evaluate the accuracy of prediction using our model.

For the purpose of modeling the streets, we note that the data collected consists of several streets which are significantly long. For example, many streets are as long as 1.5 miles. In order to capture the variation in the fuel consumption within the longer streets (due to different traffic characteristics), we divide such streets into smaller segments. Each segment is considered as one training data point for the prediction model. Note that, the collected raw data are not directly used. Instead, the model parameters (average speed, real mpg) are extracted and used for training and testing purposes.

4.2 Preliminaries

Predicting the fuel consumption using only static parameters of a street (e.g., number of traffic lights) or simple dynamic street parameters (e.g., average speed) with high accurate is challenging because there are several other factors that greatly affect the fuel consumption even if the street and the car are fixed. For example, one might drive from home to office every day at the same time and on the same route, but the fuel consumption might vary greatly with the number of traffic lights encountered that were red on a given trip (as opposed to green). Simply counting the total number of traffic lights, or even the average number of red lights encountered over a long time, does not accurately predict this time-sensitive single-trip information. Therefore, fuel consumption estimates of city traffic routes, based on any average metrics, are inherently inaccurate due to the noisy nature of the random variable being estimated. What we hope for, however, is to develop a prediction scheme whose residual error has a zero mean. In other words, if the actual fuel consumption is equally likely to be above or below the prediction, the errors will tend to cancel out (e.g., on daily commutes) and the total fuel consumption estimate over a long time will be accurate. In other words, a scheme with a zero mean error will still accurately predict one's savings at the pump, which is the basis for choosing fuel-efficient routes.

With the above in mind, we begin by plotting the variation in the fuel efficiency across various streets and cars. We plot the distribution of the miles per gallon (mpg) for the data collected for all the users in Figure 5.

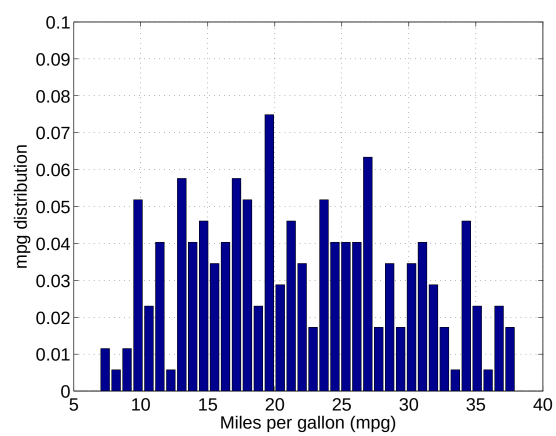


Figure 5. Figure showing the real mpg distribution for all the six users

We observe from Figure 5 that the distribution is very wide, with the mpg values varying between 5 and 40. The standard deviation of the mpg distribution is 7.75 mpg, which is pretty high. We observe that the variation in the mpg distribution is in part due to differences inherent to streets and cars and in part to time-sensitive noise added as discussed above. Hence, it is desired to have a prediction model that gives a prediction error with much smaller standard variation than the that of the overall mpg (7.75mpg).

The inputs to the prediction model include street parameters and car parameters. The goal of the prediction models is to be able to estimate the fuel consumption using only general parameters of both streets and cars which are easily acquired (e.g., number of stop signs, car make, etc). The lack of fine-grained parameters in the model makes it hard to achieve a low prediction error. However, as mentioned above, we are interested in the prediction error over a long period of time. This error can be very small if the prediction error follows a zero mean distribution. Therefore, it is reasonable to measure the accuracy of the prediction models using the sum of *signed* errors instead of absolute errors. The goal of our paper is to come up with a model with fair absolute error but with a very low signed error. Hence, we evaluate the developed model with both absolute and signed errors in upcoming sections.

With that goal in mind, we consider a simple linear model to predict the mpg for individual street segments. The mpg of streets is modeled as a function of the various street and car parameters described above. It is straightforward to compute the fuel used from the mpg as distance of the streets is known. We shall show that the model, developed in this paper, achieves an absolute error of about 11.28%, on average, while the signed error is less than 2%, which is accurate, considering that we are contemplating savings in the 10%-20% range as discussed in the feasibility section. We also show that the total prediction accuracy for long routes is much higher than the accuracy for the individual route segments, confirming the cancellation of noise.

Our linear model estimates the fuel consumption as the weighted linear combination of the parameters. The system needs to estimate the coefficient vector from the OBD-II data shared by users in order to minimize the least squared mpg error. Another advantage of the linear model is that it is possible to have a powerful online algorithm to update the coefficients of the model whenever new OBD-II data arrives, essentially using a Kalman filter.

In the rest of this section, we will incrementally build a model that achieves the aforementioned goals and results. We will begin by looking at simple models and slowly evolve into more sophisticated ones until the best is found.

4.3 The Single Car Model

In this section, we first consider models developed for Fueoogle members. These models are dedicated to their individual cars. Being fuel efficiency models of a single car, they incorporate only street parameters that the car's performance might depend on. To evaluate the accuracy of single car models in our experiments, we use data from driving the same car on one set of streets and evaluate the accuracy of model in predicting fuel consumption for a different street. In other words, the errors are computed based on the leave-one-out cross validation scheme [21].

4.3.1 Prediction with Static Street Parameters

The static street parameters under consideration include the speed limit (SL), number of stop signs (ST), and number of traffic lights (TL). We consider only static parameters in this section. Even though the error for the models presented in this section is high, we are interested in understanding the importance of the parameters that affect the fuel consumption on various streets. We use the data sets from six different cars described in Table 1.

The simplest model is a linear model that depends on only one of the parameters, SL, ST, or TL. The absolute prediction error and the corresponding signed prediction error for each car are shown in Figure 6(a) and Figure 6(b), respectively. We see from these two figures that the error in prediction is quite high, as much as 30% absolute error and 10% signed error (i.e., long-term total prediction error as a fraction of long term total consumption).

Our next step is to combine two parameters (linearly) to observe if more parameters can predict the mpg better than the single parameter models. We plot the absolute and signed errors for the three combinations of the static street parameters, (ST, SL), (SL, TL), and (ST, TL), in Figure 7(a) and Figure 7(b), respectively. We observe from these two figures that the error in prediction does not change much, which means that the mpg information contained in those static parameters are similar. We further justify this observation by considering the linear model with all three static parameters. These results are shown in Table 2.

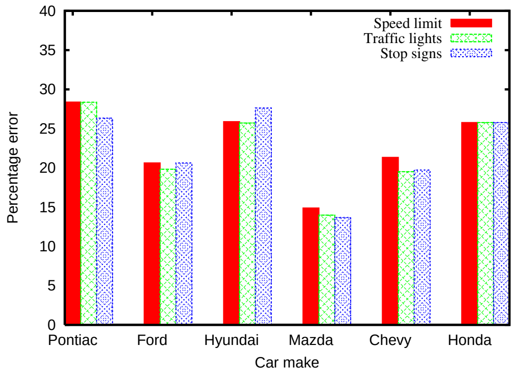
Car make	Absolute percentage error	Signed percentage error
Pontiac	27.05	9.18
Ford	22.06	6.23
Hyundai	26.65	9.08
Mazda	15.33	1.6
Chevy	18.81	6.81
Honda	28.87	1.01

Table 2. Absolute and signed prediction errors for each car/user when all the static street parameters are used in the model

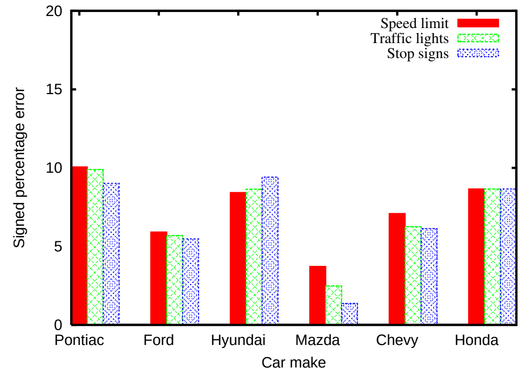
Both the absolute and the signed errors for all static parameters considered in the model are approximately the same across the simpler and more complex models. This means we have nothing to gain from using a complex model that combines the above static features. Instead, we pick the static single-parameter model that performs best (in terms of percentage signed error).

To pick that model, we compute the average error across all the users for each one parameter model in Table 3. We observe from Table 3 that the static parameter that best predicts the mpg for all the streets is the number of stop signs (admittedly, this study was performed in a small campus town).

In the rest of this section, we consider models that combine other parameters with the chosen static parameter (number of stop signs).

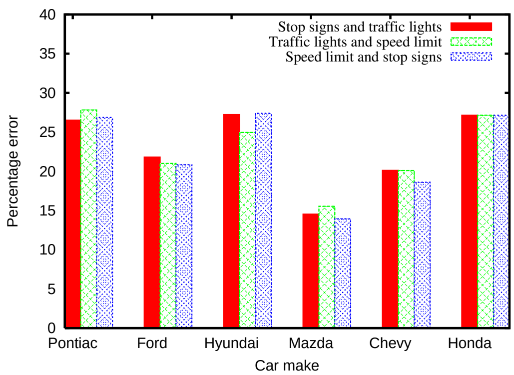


(a) Absolute prediction error

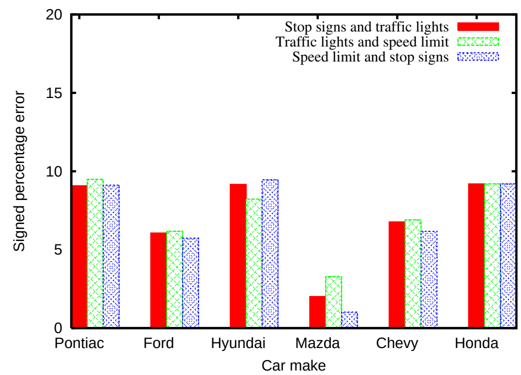


(b) Signed prediction error

Figure 6. Prediction errors for single static street parameter model for all the users



(a) Absolute prediction error



(b) Signed prediction error

Figure 7. Prediction errors for the two static street parameter model for all the users

Features	Average absolute percentage error	Average signed percentage error
Speed limit	22.82	7.31
Traffic lights	22.2	6.94
Stop signs	22.3	6.68

Table 3. Average absolute and signed prediction errors for all the cars for each of the models using static parameters only

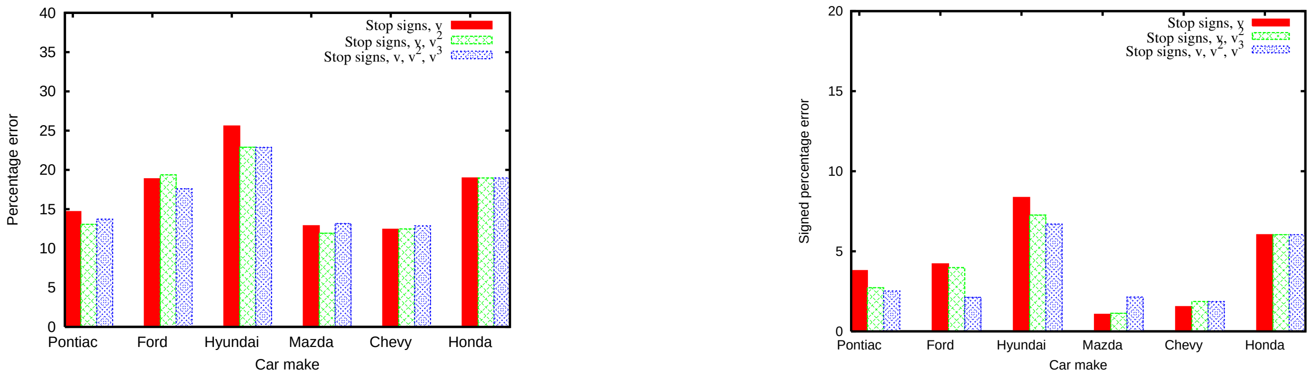
4.3.2 Prediction with Dynamic Street Parameters

The dynamic street parameters are those that vary with time. Examples of those parameters include average speed of the vehicles on the street and congestion level. First, we analyze the effect of the average speed on the mpg of the vehicle on various streets. Studies by the U.S. Department of Energy [27] show that the fuel economy (mpg) strongly relates to the average speed of vehicles. Moreover, we observe from the results in [27] that the fuel economy can be approximated by a polynomial in average speed (v) of order less than three. Now, we consider three possible models which combine the number of stop signs with v, (v, v^2) and (v, v^2, v^3). We individually train these models using data sets for different cars (to estimate coefficients of the above parameters) and evaluate them using the leave one out cross-validation scheme. The absolute prediction errors and the signed prediction errors for these three models are presented in Figure 8(a) and Figure 8(b), respectively.

We observe that both the absolute prediction errors and the signed prediction errors are significantly lower than those of the model with static parameters only, which suggests that average speed strongly correlates with fuel efficiency. However, there is no model that outperforms the others. This can be explained by the fact that the average speed and the fuel efficiency are likely to be linearly dependent in normal city traffic (the average speed is less than 40 mph). Hence we choose the best model that gives least signed prediction error. Table 4 summarizes both the absolute prediction error and the signed prediction error for the three models across all the users.

Features	Average absolute percentage error	Average signed percentage error
ST and v	17.24	4.17
ST, v and v^2	16.45	3.83
ST, v, v^2 and v^3	16.53	3.56

Table 4. Average absolute and signed prediction errors for all the cars for each of the models using number of stop signs (ST) and average speed (v)



(a) Absolute prediction error (b) Signed prediction error

Figure 8. Prediction errors for the model with stop signs and average speed (v, v^2, v^3) for all the users

The results show that the model achieving least error is the one with number of stop signs (ST), v, v^2 and v^3. In other words:

$$mpg = aST + bv + cv^2 + dv^3 + e \quad (1)$$

where a, b, c, d, and e are the coefficients derived for the vehicle in question.

Yet another dynamic factor that can improve the prediction error is the congestion level. Higher congestion levels on streets result in lower fuel efficiency, as vehicles move slowly in congested streets (thus consuming more fuel). Hence, we introduce the congestion parameter that approximates the congestion level. The congestion parameter of a certain street is defined as the ratio of the average speed of the vehicles on the street to that of the speed limit on the street. We augment the best model achieved in Equation 1 with the congestion parameter and evaluate the performance of this new model using leave one out cross validation method. The results are shown in Table 5.

Car make	Absolute percentage error	Signed percentage error
Pontiac	13.79	4.60
Ford	19.31	2.88
Hyundai	23.23	6.78
Mazda	14.36	3.05
Chevy	11.65	2.76
Honda	18.97	6.03

Table 5. Prediction error for each car/user with congestion level parameter

We compare the results of the model without the congestion parameter (Figure 8(a) and Figure 8(b)) with that of the results in Table 5. We observe that the model with the congestion parameter augmented does not improve over the model without the congestion parameters (Equation 1). This can be explained as the average speed in the original model also contains information about the congestion, thus adding a scaled version of the speed does not help in improving the model. Therefore, we choose not to add the congestion parameter into our final model.

We now consider the effect of the amount of training data used on the accuracy of prediction. We partition the data set of one car (the Pontiac Grand Am) into two sets. The first data set contains the data points recorded from one specific street. The second data set contains the rest of the data points. We train the model with the data points taken from the second data set and test the model on the first data set. We also vary the size of the training data to see the effect of the number of training data points on the performance of this model. We repeat the experiment for several training sets and average the prediction errors. The results for absolute prediction errors and signed prediction errors are shown in Figure 9(a) and Figure 9(b), respectively.

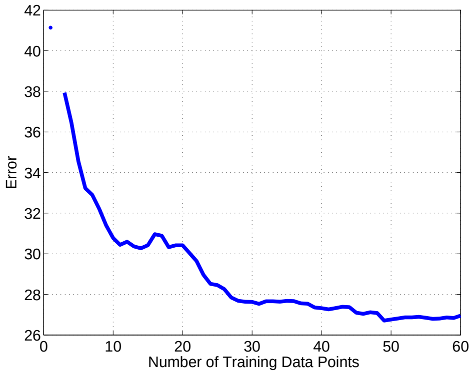
The results show that this model generalizes well to different streets even with small number of data points. On average, the model needs 40 data points to gives reasonably good prediction error for the streets that don't have any data points. This is a good number since the number of training data points for real participatory applications can be as big as thousands of data points. In the next section, we explore how models generalize across cars by incorporating car-specific parameters into the model.

4.4 The Generalized Model

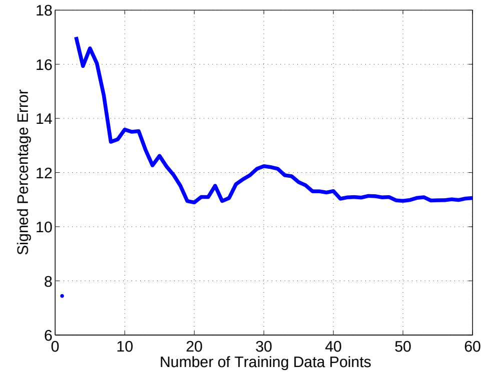
Car type is an important factor that affects the fuel efficiency. In our experiments, the fuel consumption of an SUV is higher than that of a sedan by as much as 20% on the same street at the same time of the day. Hence, it is desired for the model to be able to accurately predict the fuel efficiency across multiple types of cars. This allows Fueoogle to derive fuel efficiency of a vehicle, even when it has no prior data collected for that type of vehicle.

In order for the model to accurately predict the fuel efficiency across the different types of cars, car-specific parameters need to be incorporated into the prediction model.

The most important car-specific factor that affects the fuel efficiency is the average mpg of the specific car. Using this observation, we can use the car's average mpg as a parameter for the model. In the new model, the set of features do not change, however instead of finding the coefficients to predict the real mpg, we now find the coefficient to predict the normalized mpg, defined as the ratio of average mpg on a given

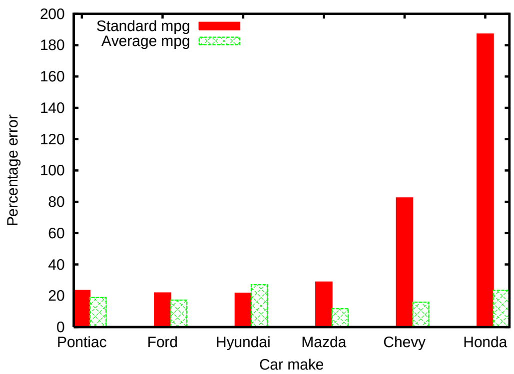


(a) Absolute prediction error

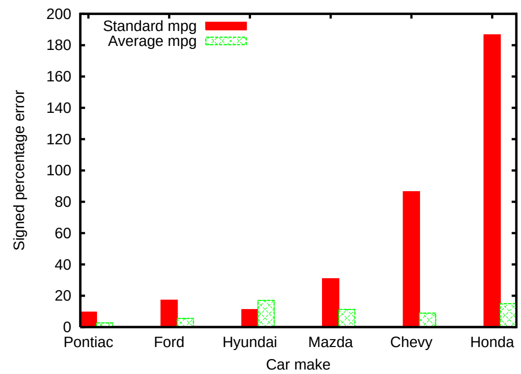


(b) Signed prediction error

Figure 9. Prediction error with increasing number of training data points for the user driving the Pontiac Grand AM, 1997



(a) Absolute prediction error



(b) Signed prediction error

Figure 10. Prediction errors when other cars' data are used for training the models that incorporate standard mpg and average mpg

street to the car's rated mpg.

The long-term average mpg for a specific car can be found in two ways: either the car owner has a sense of the average mpg or it can be provided by the manufacturer or by EPA [11]. The information provided by the manufacturer might not be a good estimate of a specific car's mpg since the average mpg of same type of car may differ as much as 10mpg [11], which may result in poor prediction performance. In this paper, we evaluate the model using both standard mpg (provided by EPA) and average mpg (provided by the owner of the car). In order to evaluate the accuracy of the model in predicting the mpg across different cars, we use all data points of one car as the testing set while use all other data points of other cars as training data. Figure 10(a) and 10(b) shows the absolute error and signed error for both models (using standard mpg and real average mpg), respectively.

Significant difference in the error performance of the model with different parameters can be seen from those figures. The model performs badly on the when using the standard average MPG from EPA which means that those values are pretty far from the accurate mpg of the car. On the other hand, the prediction model using the real average of the car performs extremely well.

In order to justify the improvement of the model after incorporating the car's mpg information, we train the single-car model described in Equation 1 across all the cars but one and test on the other car. The prediction error is then compared with the result for the multi-car model (with real car mpg). The signed prediction error for both single-car model and multi-car model is plotted in Figure 11.

We can see that the multi-car model outperforms the single-car model in most cases. It means that the normalized mpg is a better parameter for the prediction model.

Another car-specific parameter that is considered in the paper is the wheel size of the car. Fuel-efficiency will slightly drop with a smaller wheel size at a given speed because to maintain that speed, the engine and drive train have to rotate at a greater speed thus friction losses will be higher. We hypothesize that the mpg loss can be accurately linearized so it is proportional to the wheel size of the car. We evaluate our hypothesis by using wheel size of the car as a parameter to the model (which includes static parameters, dynamic parameters and car's real mpg). The model is tested using all the data points of one car and is trained using data points of

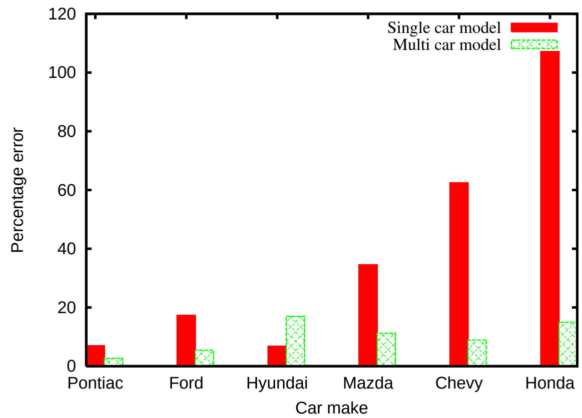


Figure 11. Signed prediction error for each car/user for the single car and multi car models

all other cars. The results is showed in Table 6.

Car make	Absolute percentage error	Signed percentage error
Pontiac	12.23	2.21
Ford	9.69	1.49
Hyundai	15.58	3.15
Mazda	13.7	2.81
Chevy	11.72	1.56
Honda	5.86	0.72

Table 6. Prediction error for each car/user when the wheel size is considered in the model

We observe from the results that the absolute error is about 11.46% while the signed error is 1.99% which is better than the previous model without wheel size information.

4.4.1 Driver Specific Model

In order to incorporate the human factor into our final model, we introduce a driver specific parameter. We want to choose a metric that captures the driving behavior of the users. For example, an individual who is used to braking hard or accelerating fast is likely to consume more fuel than a person who coasts to a stop or accelerates normally. We propose to use a parameter called the *lifetime speed variance* for a single user. This metric is the speed variance computed over the entire speed data for the given user (such a metric can be computed only for the Fueoogle members).

We compute the absolute percentage error and the signed percentage error for the six users using the method similar to the one presented in Section 4.4. These results are shown in Table 7.

We observe from Table 7 that the prediction accuracies decrease when the driver specific parameter is introduced into the final model. The average absolute percentage error is 16.53% and the average signed percentage error is 3.56%. Thus, the driver lifetime speed variance is not a useful metric for our model. Hence, our final model does not have the driver lifetime speed variance. Finally, we observe that the

Car make	Absolute percentage error	Signed percentage error
Pontiac	13.72	2.52
Honda	17.59	2.12
Chevrolet	22.87	6.7
Ford	13.16	2.14
Mazda	12.86	1.86
Hyundai	18.97	6.04

Table 7. Absolute and signed prediction errors for each car/user when the human factor is introduced into our final model

model developed in Section 4.4 achieves an average signed error of 1.99%, which is quite small. This error is acceptable for our application and hence we do not explore further parameters. We will now discuss our final model in the next Section.

4.5 Final Model Discussion

In the previous sections, we developed a linear model that can accurately predict the fuel consumption across city traffic streets and car types. We will summarize this model below. The input to the model includes:

- Static street parameters: Number of stop signs (ST)
- Dynamic street parameters: v, v^2, v^3 where v is the average vehicle speed on a specific street.
- Car specific parameters: average mpg ($\overline{mpg}$) and wheel diameter (d_w).

The final model is expressed as

$$mpg_n = aST + bv + cv^2 + dv^3 + ed_w + f \qquad (2)$$

where $a,b,c,d,e,\ f$ are the model coefficients that are estimated in the training phase. mpg_n is the normalized mpg which is computed as $mpg_n = mpg/\overline{mpg}$.

In the Fueoogle application, the static street parameters are automatically determined from existing databases such as the Red light database [16]. The average speed for each street is computed from GPS data contributed by users. For the street having no GPS information, then the average speed is guessed by the software as the average community speed. Car specific parameters are supplied by the users. The output of the prediction model is the normalized mpg for that car/street. Fueoogle multiplies this number with the vehicle average mpg to get the real mpg for that car/street.

We now evaluate the overall performance of the final model using the leave one out cross validation scheme on the all the data set of six cars. As discussed in Section 4.2, it is desirable for the standard deviation of the error of the final model to be smaller than the standard deviation of the data itself. In addition, we are only concerned about the signed error distribution since it represent the typical error behavior when estimating long street segments. The error distribution for signed error of the final model is plotted in Figure 12.

The error distribution in Figure 12 resembles a Gaussian distribution with zero mean and standard deviation of 0.1434 mpg, which is significantly smaller than the standard deviation of the real data itself (7.75 mpg). Two sigma rule tells us

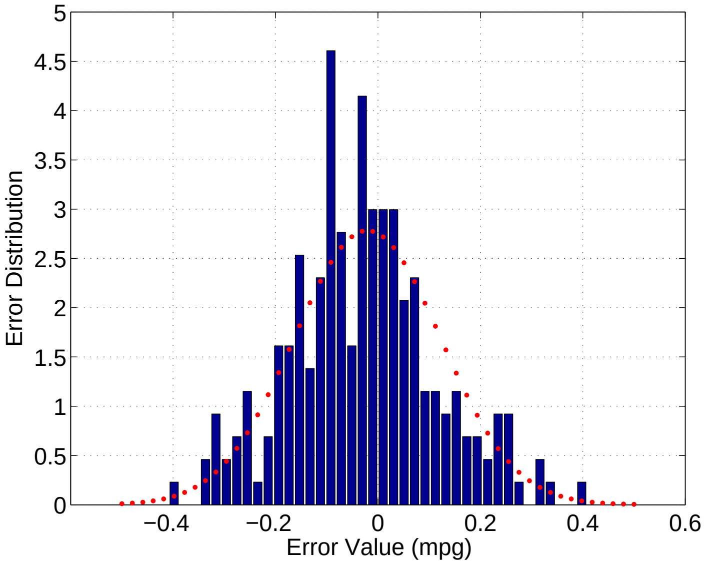


Figure 12. Distribution of the signed error value for the general model for all the cars and streets

that the prediction error is less than 0.2868 mpg with probability of 95% which is very good.

4.6 Updating the Model

Since Fueoogle is a participatory sensing application that provides a long term service for the community, there is a need to update the model using new OBD-II data to prevent inaccuracy due to out-dated model. To update the prediction model, one solution is to augment old data with newer data and solve the least squared optimization to find new model coefficients. However this solution does not scale because of the accumulation of OBD-II data over years and the computational complexity to solve the least squared optimization with huge input data. Therefore there is the need to find an online algorithm to update the model coefficients using just a small number of past variables and new OBD-II data. In this paper, we present an online algorithm to update the model coefficients based on incremental gradient method [6].

For simplicity, we denote x as the current model coefficient vector with m elements corresponding to m features discussed in Section 4.5. The set of new OBD-II data features is denoted as as $C = (C_1, C_2, \ldots, C_n)$, and $Z = (z_1, z_2, \ldots, z_n)$ is the target MPG get from the new OBD-II training data. We compute the model coefficient by minimizing the following unconstrained quadratic optimization

$$x_i = argmin \sum_{j=1}^{i} |z_j - C_j x|^2 \tag{3}$$

The incremental gradient method iteratively finds the the optimal coefficient vector x for each $1 \leq i \leq n$. The optimal solution of x at step i is denoted as x_i. The incremental gradient solution for this linear curve fitting as follow

$$x_i = x_{i-1} + H_i^{-1} C_i^T (z_i - C_i x_{i-1}) \tag{4}$$

H_i is also iteratively computed using following equation

$$H_i = H_{i-1} + C_i^T C_i \tag{5}$$

The initial condition is $x_0 = 0$ and H_0 being arbitrary positive definite matrix. Interestingly, this equation is a realization of the Kalman filter [6]. Readers are encouraged to refer to [6] for more discussion of the incremental gradient method for linear system.

In order to update the coefficient, we only need to store an *mxm* matrix H in addition to the old coefficient which is extremely resource efficient. One property of the Kalman filter is that the model parameters converge to a new state very fast when the characteristic of the system change. This guarantees our system to be up to date when there are changes in traffic characteristics.

5 Evaluation

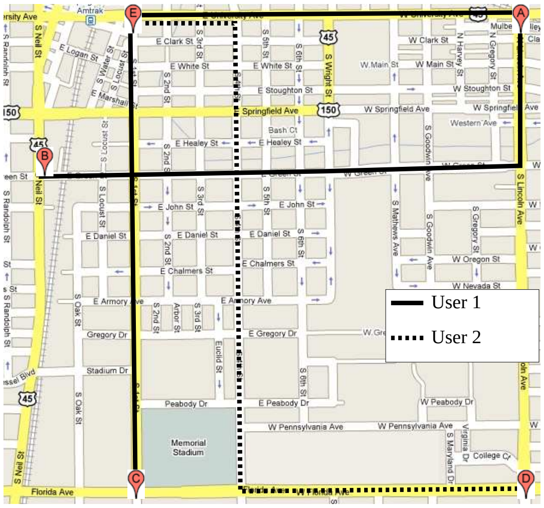
In a sense, the performance of the fuel consumption models we presented has already been evaluated in the context of deriving the best model structure. We therefore present in this section only a small number of additional experiments that confirm the efficacy of the winning model. We evaluate how the system performs both in terms of the accuracy of the model in predicting end-to-end fuel consumption for long routes as well as in terms of ability to find the most fuel efficient paths. We use the data collected from the second set of experiments (the data collection is described in Section 4.1) to compute these results.

For each of the considered routes, we compute the actual fuel consumed and the predicted fuel consumption from the Fueoogle system (which uses the final model described in the previous Section). Figure 13 shows the routes of two different users. These computed results along with the percentage error for the end-to-end path are shown in Table 8.

Car make/ Path	**Actual fuel (gallons)**	**Predicted fuel (gallons)**	**Percentage error**
Pontiac (A-B)	0.0767	0.0757	1.3
Pontiac (A-C)	0.0786	0.0760	3.3
Ford (D-E)	0.0980	0.0948	3.2
Chevy (A-B)*	0.0817	0.0897	2.3
Chevy (A-B)**	0.0789	0.0857	8.6

Table 8. Table showing the actual and predicted fuel consumption in gallons for the routes shown in Figure 13 along with the percentage error in prediction. The first entry for Chevy (A-B)* is for the fastest route and the second entry (A-B)** is the most fuel efficient route chosen by Fueoogle.

We observe from Table 8 that the percentage errors for the end-to-end routes for most of the paths are quite small. The average prediction error for the end-to-end path is 3.75%. This demonstrates that Fueoogle achieves a small percentage error in predicting the fuel consumption for the end-to-end paths.



(a) Figure showing the routes chosen for two different users for evaluating the goodness of our final model. User 1 drives a Pontiac Grand AM, 1997 and User 2 drives a Ford Taurus 2001.

(b) Figure showing two routes chosen for one user between two landmarks. The most fuel efficient route which Fueoogle picks is marked by a solid line, whereas the fastest route is marked by dashed lines. The user drives a Chevrolet, Prizm, 1998.

Figure 13. Maps showing the driving routes for the purpose of evaluation

Further, we asked our participants to choose a random route between two points and measure fuel consumption, then ask Fueoogle and follow its directions between the same endpoints, measuring fuel consumption again. All participants reported fuel savings. One such experiment is shown in Figure 13(b). The user drove on two different paths between landmarks A and B. One was the fastest route from MapQuest and the other is a route by Fueoogle. Fueoogle picked the route that consumed less fuel, when compared to the fastest route from MapQuest. The most fuel-efficient route is marked by a solid line in Figure 13(b).

We conclude from the above observations that Fueoogle predicts the fuel consumption for end-to-end routes with a high accuracy and also chooses the most fuel efficient route.

6 Related Work

We divide this section into three parts, the first part presents related work in participatory sensing and the second examines fuel efficiency related literature.

6.1 Participatory Sensing

The concept of participatory sensing was introduced in [8]; participatory sensing is where individuals are tasked with data collection which is then shared for a common purpose. A broad overview of such applications was later provided in [1]. Several early applications have been published. Examples include CenWits [18], a participatory sensing network to search and rescue hikers, CarTel [19], a vehicular sensor network for traffic monitoring, BikeNet [10], a bikers sensor network for monitoring popular cyclist routes, and ImageScape [23], cellphone camera networks for sharing diet related images. Our application, Fueoogle, introduces a novel participatory sensing application that enables individuals to obtain fuel efficient routes within a city.

6.2 Fuel Efficiency

A comprehensive study that provides optimal route choices for lowest fuel consumption is presented in [12]. In the paper, fuel consumption measurements are made through the extensive deployment of sensing devices (different from the OBD-II) in experimental cars. These fuel consumption measurements are then used to compute the lowest fuel consumption route. As opposed to the work in [12], our paper uses a sparse deployment to build mathematical models for predicting fuel consumption on streets that lack the real measurements. In [7], the influence of driving patterns of a community on the exhaust emissions and fuel consumption were studied. Feedback was provided to the community regarding the driving patterns to cut back on the fuel consumption and exhaust. A driver support tool, FEST, was developed in [9]. FEST uses sensors installed in the car along with a software to determine the driving behavior of the driver and provide real-time feedback to the individual for the purpose of reduction in fuel consumption. An extension to FEST that includes more experiments and further evaluation can be found in [28]. A feedback control algorithm was developed in [24] that determines speed of automobiles on highways with varying terrain which achieve minimal fuel consumption. An extension to the work in [24] was developed in [17]. In [17], suggestions of driving style to minimize fuel consumption were made for varying road and trip types (e.g. constant grade road, hilly road). The problem was formulated using a control theoretic approach.

In a separate study [20], it was shown that rising obesity has a significant impact on the total fuel consumption of the US. Models were developed that studied the impact of obesity on the amount of fuel consumed in passenger vehicles.

In our work, we develop models that estimate fuel consumption of streets based on measured parameters of the given street (e.g. speed limit of street, number of traffic lights). These estimates of fuel consumption of streets are then used to compute fuel-optimal routes.

7 Conclusions and Future Work

In this paper, we developed a participatory sensing application, called Fueoogle, that provides a service which computes fuel efficient routes from one point to another in the city where the authors reside. This service relies on OBD-II data collected by a set of users who share their data with Fueoogle using a previously published participatory sensing framework, called PoolView. The paper shows that significant fuel savings can be achieved by using Fueoogle in a larger community, which not only reduces the amount of money spent by people on their daily gasoline consumption, but also has a positive impact on the environment by reducing the amount of CO_2 emissions. We show that Fueoogle can utilize a sparse deployment to estimate the fuel consumption on streets that lack OBD-II measurements, as well as estimate fuel consumption of vehicles using data on other vehicles. Fueoogle achieves this by using the model developed in this paper to estimate the fuel consumption for those streets and vehicles that lack their own OBD-II measurements. Our future work will address the impact of privacy-preservation mechanisms such as data perturbation on the correctness of aggregate fuel statistics computed by Fueoogle, as well as gain experience from a long-term deployment.

8 References

[1] T. Abdelzaher et al. Mobiscopes for human spaces. *IEEE Pervasive Computing*, 6(2):20–29, 2007.

[2] Actron. Elite autoscanner. http://www.actron.com/product_category.php?id=249.

[3] Auterra. Dashdyno. http://www.auterraweb.com/dashdynoseries.html.

[4] AutoTap. Autotap reader. http://www.autotap.com/products.asp.

[5] AutoXRay. Ez-scan. http://www.autoxray.com/product_category.php?id=338.

[6] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, second edition, 2003.

[7] K. Brundell-Freij and E. Ericsson. Influence of street characteristics, driver category and car performance on urban driving patterns. *Transportation Research, Part D*, 10(3):213–229, 2005.

[8] J. Burke et al. Participatory sensing. Workshop on World-Sensor-Web, co-located with ACM SenSys, 2006.

[9] V. der Voort. Fest - a new driver support tool that reduces fuel consumption and emissions. *IEE Conference Publication*, 483:90–93, 2001.

[10] S. B. Eisenman et al. The bikenet mobile sensing system for cyclist experience mapping. In *Proc. of SenSys*, November 2007.

[11] EPA. Emission facts: Greenhouse gas emissions from a typical passenger vehicle. http://www.epa.gov/OMS/climate/420f05004.htm.

[12] E. Ericsson, H. Larsson, and K. Brundell-Freij. Optimizing route choice for lowest fuel consumption - potential effects of a new driver support tool. *Transportation Research, Part C*, 14(6):369–383, 2006.

[13] R. K. Ganti, N. Pham, Y.-E. Tsai, and T. F. Abdelzaher. Poolview: Stream privacy for grassroots participatory sensing. In *Proc. of SenSys '08*, pages 281–294, 2008.

[14] Garmin eTrex Legend. www8.garmin.com/products/etrexlegend.

[15] Google. Google maps. http://maps.google.com.

[16] GPS POI. Red light database. http://www.gps-poi-us.com/.

[17] J. N. Hooker. Optimal driving for single-vehicle fuel economy. *Transportation Research, Part A*, 22A(3):183–201, 1988.

[18] J.-H. Huang, S. Amjad, and S. Mishra. Cenwits: a sensor-based loosely coupled search and rescue system using witnesses. In *Proc. of SenSys*, pages 180–191, 2005.

[19] B. Hull et al. Cartel: a distributed mobile sensor computing system. In *Proc. of SenSys*, pages 125–138, 2006.

[20] S. H. Jacobson and L. A. McLay. The economic impact of obesity on automobile fuel consumption. *Engineering Economist*, 51(4):307–323, 2006.

[21] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proc. of International Joint Conference on Artificial Intelligence 2 (12)*, pages 1137–1143, 1995.

[22] MapQuest. http://www.mapquest.com.

[23] S. Reddy et al. Image browsing, processing, and clustering for participatory sensing: Lessons from a dietsense prototype. In *Proc of EmNets*, pages 13–17, 2007.

[24] A. B. Schwarzkopf and R. B. Leipnik. Control of highway vehicles for minimum fuel consumption over varying terrain. *Transportation Research*, 11(4):279–286, 1977.

[25] Traffic. Real-time traffic conditions. http://www.traffic.com/.

[26] US Census Bureau. Tiger shapefiles. http://www.census.gov/geo/www/tiger/.

[27] U.S. Department of Energy. Driving more efficiently. http://www.fueleconomy.gov/feg/driveHabits.shtml.

[28] M. van der Voort, M. S. Dougherty, and M. van Maarseveen. A prototype fuel-efficiency support tool. *Transportation Research, Part C*, 9(4):279–296, 2001.