PREDICTION OF FUEL ANALYSIS USING URBAN BIG DATA

R. Logeshwari
Assistant Professor, Department of Computer Science and
Engineering, SRM Institute of Science and Technology
Chennai, India

Divyansh Gautam
Student, Department of Computer Science and Engineering,
SRM Institute of Science and Technology
Chennai, India

Aniruddha Joshi
Student, Department of Computer Science and Engineering,
SRM Institute of Science and Technology
Chennai, India

*Abstract:*Ability to model and predict the fuel consumption is vital in enhancing fuel economy of vehicles in transport management. There are several internal factors such as distance, load and vehicle characteristics, as well as external factors such as road conditions, traffic, and weather on which fuel consumption of a vehicle is dependent. However, not all these factors may be available or measured for the fuel consumption. Providing real time traffic information in metropolitan cities is desired since it not only helps to manage the traffic management but also save the time of travelers and reduces the vehicle fuel consumption. To obtain the traffic information from number of sensors on every road segments or intersections is difficult due to large number of installations. Getting the accurate information of current and near term future traffic flows of different road links in a traffic network has a wide range of applications which includes the forecasting of the traffic flow, navigation of vehicles and traffic congestion management. We considered a case where only subset of three factors is easily available which are vehicle characteristics, traffic dataset and road distance. Hence, the challenge is to model and/or predict the fuel consumption only with available data, and also taking as much as influence from other internal and external factors. Machine Learning (ML) is suitable in such analysis, as the model can be developed by learning the patterns in the available data. In this paper, we use algorithm that is used in google maps such as Gaussian Naïve Bayes and Page Rank which provide the different routes and methods like image processing of maps to extract the different RGB values of different routes which helps to predict the fuel consumption of the vehicle. Finally, after predicting the fuel consumption for different paths, the best path gets generated in terms of less fuel consumptions.

*Keywords:*VehicleCharacterstics, Fuel Consumption, Traffic Dataset, Image Processing, RGB Values

I. INTRODUCTION

The study evaluates the method of statistical analysis for predicting the fuel consumption using urban big data. The idea is to use the historical and available data which describes the traffic situations, road networks and vehicle characteristics to predict the fuel consumption. The current system provides the information about the time and distance for travelling from one point to another which also gives the real time traffic information. The proposed idea is to use the road distance and traffic information and the vehicle mileage to provide the fuel consumption based on the starting and ending point for the vehicle. The general problem description is to examine the large number of attributes which gives the idea for fuel consumption and using of the algorithms to find a solution from the attributes. Attributes includes the vehicle characteristics, traffic flow and road distance. Research has been made into how to do such predictions using the available data. The goal of the companion project is to develop a real time application to predict the fuel consumption taking into historical and real time information of traffic, and collecting the vehicle characteristics based on the engine CC and their mileage and provide the best path based on less fuel consumption. Earlier research has been done to find out the different traffic jam analysis based on trajectory data using GPS trajectories [1]. One goal of the study is to use different algorithm approaches for prediction from the set of descriptive attributes. Study of different description of traffics sensory data which includes collection of GPS reports and characteristics of sensory data has been analyzed earlier which provides some information regarding traffic condition[2]. Study of the

attributes included the vehicle mileages, different road networks for two points having different distances, real time traffic flows for different time period on a day. In a city, the fuel consumption varies based on the real time traffic flows, so straight forward traffic is related to the time taken for reaching one place to another. If the traffic is more the time taken to reach the destination will be more which leads to the more fuel consumption. To analyze the traffic flows, first we have to analyze the average speed of different vehicles at different time period. It means for a particular place, what will be the average speed of a vehicle in morning, afternoon, evening and night likewise because the speed of any vehicle changes with the traffic. By collecting the different speeds for different routes or paths of a vehicle for different time period for 1 or 2 weeks will give us the approximate average speed of the vehicle on which it travels mostly on different traffic conditions. Google maps will provide the time taken by a vehicle travelling from start position to destination position based on the traffic flow. For different time variation the, travelling time will be analyzed based on traffic flow. Fuel consumption is usually estimated based on distance and the mileage of vehicle. It is calculated as division of two factors that is total distance and the mileage of the vehicle. However, there are different types of uncertainty will be introduced, when the mileage of a vehicle is measured in different ways. The mileage of a vehicle varies with the region and road conditions. Also the datasets have been created using the google maps which shows the real time traffic based on different colors. It helps to analyze the traffic flow, which helps to get the average speed for the particular path. So, with the help of speed and time, the fuel consumption could be calculated for the different path for the particular starting and

destination point. The main contributions of this paper includes that we identify the different kind of datasets based the real time traffic, different routes and paths and vehicle mileages based on that fuel consumption get calculated. There is no existing model for predicting the fuel consumption in vehicles. The novelty of idea is the study of collected data sets and connecting to predict the fuel consumption. The rest of the paper is organized as follows. We first present the related work. Then we define the proposed work and describe the framework of our system. Then methodology and experimental evaluation. Finally, we conclude our work.

II. RELATED WORK

Our related work starts with the traffic optimization. With the use of google maps, it helps us to get more urban traffic data. And with help of these data, a research on traffic have been done which includes the study of different colors based on the traffic, traffic congestions at different time period in a day, and study of typical traffic. With the help of all these, it helps us to create the traffic datasets for getting the best path. From the google maps, typical traffic terms has been analyzing at different time stamps. Data sets are created having the travelling time between source and destination for different time stamps of the day. Next is time based approaches, much of previous work provides the travelling time only based on the traffic flow to reach destination. Calculating the fuel consumption only includes the limited factor like distance and mileage only, but traffic is also an important factor which affects the fuel consumption of the vehicle. Google maps provide us the minimum and maximum travelling time for different time stamps, so the data is stored and analyzed to generate some result sets for calculating the fuel consumption. And last is urban computing, it is described in detail in paper[3], with the increasing number of vehicles and traffic data, researches on urban computing become more attractive. Analyzing the vehicle characteristics based on different cities and different traffic conditions may varies, which affects the mileage of the vehicles. The datasets are created based on an average mileage a vehicle gives. Live tracking of the vehicle and getting the mileage based on that is difficult which requires some hardware sensors. So, on an average vehicle datasets are being created and also the data sets are created based on the average speed of vehicles at different timings at different traffic situations, so that by analyzing these all factors helps us to calculate the fuel consumption more accurately.Some related work is also done for getting the vehicle characteristics. Vehicle mileage dataset is created for two wheeler's which is having the vehicle names and there mileages.

III. PROPOSED WORK

A. Preliminary

In this paper we propose to analysis the best efficient route based on the fuel consumption of a particular vehicle. Monzon[4] has given the effects of traffic congestions which affects the fuel consumption. There are many factors in prediction of the fuel efficiency of a trip but in this paper we are using only a handful of resources which are easily available namely the Source and Destination, the Traffic, the Vehicle of the user. Source and Destination are provided by the user and the distance between them, the time of travel and the traffic are readily used by us from Google Maps. The Vehicle type is selected from our database by the user which has predefined vehicles and their mileages in the database. We in this paper are trying to use the mileage, traffic and the time

taken for the trip to predict the efficient route. The process of analysis goes as follows:

Firstly, the user enters/selects the Source and Destination. He then selects his Vehicle, upon which the data is loaded in the application from the database about that entered vehicle, source and destination. The source and destination points are given to the Map (API) which provides us with the suggested paths for travel. These paths will have their distance and travel time which is given to the application for analysis. Along with this, the application also gets the typical (or live) traffic on each path for that particular set of source and destination points. This is also used for calculation of average speed in the traffic. Earlier Jacek Oskarbski[5] has given the idea to calculate the average speed based on traffic system data.

The Google Maps traffic layout indicates traffic density through colors. It has different colors for different traffic densities namely, Dark Red for Heavy Traffic, Red for Moderate Traffic, Yellow for Low Traffic and Blue if the road is conveniently empty for travel. According to our personal work done on various bikes, we have found out that vehicles in these color zone travel between a particular speed range. So we took and set an average of that speed range for each color. The images of these typical (or live) traffic are processed and by taking their particular RGB values, the percentage of these colors in the image is calculated. These percentages can tell us the distance covered by particular traffic density type in the whole trip from the total distance of the trip. Different segments (distance) of the whole route are travelled with different speed, so we can calculate the average speed of the total trip at that particular timestamp using the following average speed formula.

$$\text{Average Speed} = \frac{\text{Total Distance}}{\{\frac{d1}{s1} + \frac{d2}{s2} + \frac{d3}{s3} + \frac{d4}{s4}\}} \quad (1)$$

Where, d1 and s1 are the distance and avg. speed of the Dark Red Segment.
d2 and s2 are the distance and avg. speed of the Red Segment.
d3 and s3 are the distance and avg. speed of the Yellow Segment.
d4 and s4 are the distance and avg. speed of the Blue Segment.

Fuel Consumption of a particular trip is calculated based upon the distance of the trip and the riding vehicle's mileage which is given by the formula:

$$\text{Fuel Consumption} = \frac{\text{Total Distance}}{\text{Mileage of the Vehicle}} \quad (2)$$

Here, the Total distance can be expressed in terms of the Average Speed of the trip and the Total Time taken for travel. So the above equation could be rewritten as,

$$\text{Fuel Consumption} = \frac{\text{Average Speed} \times \text{Total Time}}{\text{Mileage of the Vehicle}} (3)$$

Now when we have all the parameters like Average Speed (calculated above), Total Time (from the Map) at that timestamp and the Mileage (pre-stored in the Database), we could calculate the approx.
Fuel Consumption for that particular route and also for the other suggested substitute route for that particular pair of source and destination.

Among all these routes, the best route is selected which uses the least fuel for that particular vehicle for that journey between source and destination at that particular time and themost Fuel Efficient Route is predicted and suggested to the user as a result set for his preferences.

B. Framework

Our framework describes the two access control modules, one is for admin part and another is for traveler part. Traveler part is UI for the traveler from which the user going to check out the path. Datasets are created and saved in admin part which includes the datasets of different locations with their different path distance, and datasets of vehicle, and datasets of travel time and average speed at different time stamp. The user needs to enter the source and destination and have to select the vehicle name which is already stored in the admin database. When a user check out the path, various paths will be provided by admin part and at the same time fuel consumption calculation has been calculated for the given path. Next, the user will checkout the fuel consumption for various paths and among them the best efficient path in terms of fuel consumption will get generated.
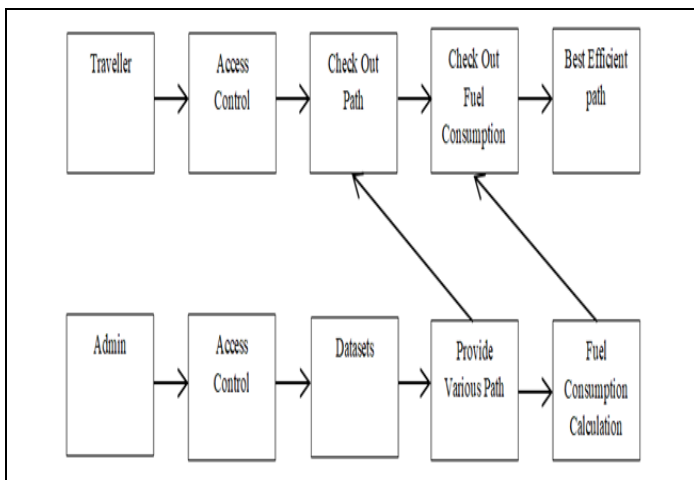

Figure 1. Framework of the system

IV. EXPERIMENTS

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

A. Datasets

In the experiments, we validate our prediction system using the three kinds of real world datasets.

- Vehicle Mileage: The vehicle mileage dataset is generated by about over 50 two wheeler vehicles. We use these datasets to calculate the fuel consumption. Based on the mileage only the fuel consumption varies, so user need to select their vehicle from the dataset and based on that the fuel consumption get calculated.

Table I. Sample Vehicle Mileage Dataset

Sr. No.	Vehicle Name	Mileage
1	Honda Activa 4G	60 km/l
2	Bajaj Pulsar	65 km/l
3	Yamaha YZF R15	47 km/l
4	TVS Jupiter	62 km/l
5	Bajaj Dominar 400 2018	26.5 km/l

- Road Distance: Road distance dataset is generated by getting the various distance of different path for same source and destination (i.e. for some source s and destination d, there are P1,P2….Pn paths) from google maps. With the paths the average speed of vehicle for time stamp is been calculated and that also get stored in same dataset with their respective path.

Table II. Sample Road Distance Dataset

Sr.No.	Source	Destination	P1	P2	P3
1	Adyar	Nungambakkam	9.6 km	9.5 km	9.6 km
2	OMR	Tambaram	23.5 km	24.3 km	26.9 km
3	Ashok Nagar	Koyembedu	6.3 km	6.2 km	6.9 km
4	Velachery	Koyembedu	15 km	15.1 km	18.1 km
5	Tambaram	KK Nagar	19.1 km	28.2 km	33.8 km

- Traffic: The traffic dataset is going toused as statitcal data in terms of time. Google maps provide the travelling time which is directly related to the traffic, so instead of taking traffic, travel time is taken. So, at different time stamp, the average travel time dataset has been created for different locations for their different paths, which helps us to relate the traffic in terms of time.

B. Image Processing

In Image Processing, the Google Maps traffic layout indicates traffic density through colors. LinhuiLi[6] proposed the methodology on traffic scene segmentation based on RGB image values. Google maps has different colors for different traffic densities namely, Dark Red for Heavy Traffic, Red for Moderate Traffic, Yellow for Low Traffic and Blue if the road is conveniently empty for travel. For different routes and paths for different time stamp, the images have been captured from google maps. These images includes the typical traffic with their respective colors. The images of these typical (or live) traffic are processed and by taking their particular RGB values, the percentage of these colors in the image is calculated. These percentages can tell us the distance covered by particular traffic density type in the whole trip from the total distance of the trip. By getting the distance, the average speed of the vehicle for the particular path at different time stamp is been calculated accordingly.Average speed of a vehicle is calculated for different traffic colors, like for dark red the average speed of a vehicle is 7km/h, likewise for all the colors the average speed of vehicle is calculated according to traffic

conditions. This help us to calculate on an average speed of he vehicle for the whole path including the traffic. By image processing, we have calculated the average speed which helps to calculate the fuel consumption of the vehicle.

Table III. Different RGB Values

Sr.No.	Traffic Colors	R	G	B	Average Speed
1	Dark Red	190	0	0	7 km/h
2	Red	255	0	0	15 km/h
3	Yellow	255	158	0	28 km/h
4	Blue	0	179	253	43 km/h

C. Result

By taking the percentage and calculating the traffic from the four colours, the average speed is calculated at particular time stamp. Then the fuel consumption is calculated by using the average speed, travelling time and mileage of vehicle which generates a result set for the particular path, likewise for other paths the result set is been generated for particular source and destination and among them which one having less fuel consumption get selected as an efficient path in terms of fuel consumption.

Table IV. Sample ComarisionResult Set

Path	Distance	Travel Time	Average Speed	Mileage	Result by"(2)"	Result by"(3)"
P1	8.9 km	30 min	29.4 km/h	60 km/l	148.3 ml	245.0 ml
P2	9.3 km	32 min	33.5 km/h	60 km/l	155.0 ml	297.7 ml
P3	6.9 km	21 min	34.7 km/h	60 km/l	115.0 ml	202.4 ml

V. CONCLUSION

In this paper, from the perspective of big data, we predict the fuel consumption based on three real world datasets which is traffic flow, road networks and vehicle characteristics. We identified the datasets and based on the relationship between them fuel consumption is predicted. Image processing help us to calculate the real time traffic distance and which helps to

calculate the average speeds of vehicles on the particular path with respect to traffic. Datasets are created which contains the different average speeds of vehicles at different time stamp which help to predict the fuel consumption. Also google maps provide the different paths and routes with the respective travel time. Finally, a mobile application is created based on all these aspects and whenever a user opens the application, the user needs to enter the source and destination and there vehicle type. Based on all the inputs by the user, our application calculates the fuel consumption on real time basis. Accordingly, we have used some mathematical formulas for providing the fuel consumption based on that providing the best path for travelling with minimum fuel consumption. In the future, we will extend this work in several directions, like for all type of vehicles and for more cities, in order to better predict the fuel consumption. Subsequently we would try to capture the real time traffic other than real time and also to use Machine learning techniques like neural network to generate the maximum accuracy for prediction and to we try to get more and more data and relate them for the prediction. Also we try to implement our project in automobile industries, cab taxi services, transportation logistics etc.

VI. REFERENCES

[1] Z. Wang, M. Lu, X. Yuan, J. Zhang, and H. Van De Wetering, "Visual traffic jam analysis based on trajectory data," IEEE Transactions on Visualization and ComputerGraphics (Volume: 19, Issue: 12, Dec. 2013).

[2] K. Ota, M. Dong, H. Zhu, S. Chang, and X. Shen, "Traffic information prediction in urban vehicular networks: A correlation based approach," in Wireless Communications and Networking Conference (WCNC), 2011 IEEE. IEEE, 2011, pp. 1021–1025.

[3] Y. Zheng, L. Capra, O. Wolfson, and H. Yang, "Urban computing:concepts, methodologies, and applications," ACM Transactions onIntelligent Systems and Technology (TIST), vol. 5, no. 3, p. 38, 2014.

[4] Monzon, Andres. "Measuring the effects of traffic congestion on fuel consumption 2",unpublished.

[5] Jacek Oskarbski,Krystian Birr, Michał Miszewski, Karol Żarski, "Estimating the average speed of public transport vehicles based on traffic control system data," Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2015 International Conference.

[6] Linhui Li, Bo Qian, Jing Lian, Weina Zheng, Yafu Zhou, "Traffic Scene Segmentation Based on RGB-D Image and Deep Learning," IEEE Transactions on Intelligent Transportation Systems (Volume: PP, Issue: 99).