A novel approach for vehicle specific road/traffic congestion

Student Name: Shilpa Garg

IIIT-D-MTech-CS-14-MT12049
March 22, 2014

Indraprastha Institute of Information Technology
New Delhi

<u>Thesis Committee</u>
Dr. Pushpendra Singh (Chair)
Dr. Vikram Goyal(Internal)
Prof. Parmeshwaram Ramanathan(External)

Submitted in partial fulfillment of the requirements
for the Degree of M.Tech. in Computer Science,

Keywords: Ubiquitous Computing; Vehicle Classification; traffic congestion; AdaBoost;

Certificate

This is to certify that the thesis titled **"A novel approach for vehicle specific road/traffic congestion"** submitted by **Shilpa Garg** for the partial fulfillment of the requirements for the degree of *Master of Technology* in *Computer Science & Engineering* is a record of the bonafide work carried out by her under my guidance and supervision in the Mobile and Ubiquitous Computing group at Indraprastha Institute of Information Technology, Delhi. This work has not been submitted anywhere else for the reward of any other degree.

Dr. Pushpendra Singh
Indraprastha Institute of Information Technology, New Delhi

Abstract

Large population, poor road infrastructure, and rapidly growing economies lead to severe traffic congestion in many parts of the world. The problem is exacerbated by increased diversity in vehicle types and poor adherence to lane discipline. Existing approaches for detecting traffic congestion do not deal with the diversity of vehicle types and lack of lane discipline. We propose a novel approach to detect congestion levels as per on the vehicle type. We present a threshold-based heuristic and an improved classification algorithm to infer the main regions of high traffic in a city. We do transportation mode classification from the real time crowd-sourced smartphone sensor data followed by congestion level detection. Our approach can be useful in congestion management by suggesting optimal alternate route as per the vehicle type, thus saving travel time and reducing fuel consumption and emission reduction. The proposed congestion detection system was able to predict more than 80% major hot-spots of traffic classified as per different vehicle types.

Acknowledgments

Foremost, I would like to express my sincere gratitude to my advisor Dr.Pushpendra Singh for the continuous support of my Masters study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Masters study.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr.Vikram Goyal and Prof. Parmesh Ramanathan for their encouragement, insightful comments, and hard questions.

I wish to thank many people for their direct or indirect help in achieving this goal. Special thanks to my parents for their support and encouragement throughout my study, their faith in me ignites that spark to do anything for their happiness.Last but not least I would like to thank all my friends for being with me at each step when I need their support. This thesis would never be successful without your support and love.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Travelling is one of the daily routine activities for most people. Commuters use different type of vehicles, car bus, motorbike etc., every day for reaching their destination. Roads and vehicular traffic are essential part of the day-to-day life of the people. Hence, analyzing and monitoring the traffic patterns and problems concerning the same is of prime concern [10].

Prior work is focused on the road and traffic condition monitoring of developed countries, where there is lane based traffic and quality of roads is generally good. It provides a uniform flow of the traffic on roads. This can be analyzed and monitored using the characteristics like volume & density of the traffic, and average speed of the vehicles [15], [9], [3]. This monitoring is employed in Intelligent Traffic System (ITS) [1], [2] which are developed using different hardware installations like loop vehicle detectors, cameras at intersection junctions, accelerometers inside vehicles [22], [11]. These approaches are highly cumbersome, incur high installation and maintenance cost and require large manual effort which can induce error while deployment or collecting the data.

In emerging economies like India, the traffic flow is more chaotic as compared to developed countries. The condition is contrasted and impacted due to many factors like heterogeneous vehicles types, socio-economic reasons, irregular driving patterns, poor quality of roads and liberal honking [22], [18], [6]. Monitoring road and traffic conditions equipped with such variations poses many challenges. These scenarios though require intense analysis and are complex but at the same time provide rich information about the prospects of traffic monitoring. Traffic congestion management has diverse set of applications like accident detection, optimal route mapping, user profiling, saving fuel consumption, and emergency services, road quality map, and checking driving tendency [18], [14], [19]. In India, vehicular population ratio of 2-wheelers and 3 wheelers is more than that of developed countries. One important factor contributing in traffic congestion is surface area of roads and vehicle density. The average speed and road surface occupancy vary from vehicle to vehicle. With various types of vehicles on the road, with different dimensions and need, it is of prime importance and essence that the traffic monitoring and management should be vehicle oriented. Vehicle classification offers insight to solution of providing congestion details. It pillars the fact that traffic congestion level may not be similar to each vehicle type for example what might be a heavy traffic for a large vehicle, it is not so for a light vehicle like motorbike. Here, we advocate the traffic congestion detection based on vehicle classification.

To address this problem, we present a vehicle specific novel approach to detect traffic congestion levels on Indian roads using the Smartphone. Our approach detects the city traffic congestion and hence enables a new direction for real world applications development for traffic anomalies. We use the proven methods for vehicle classification followed by the congestion level detection based on vehicle. Our approach involves the Smartphone which avoids the expensive infrastructure and dedicated sensors. [24]. Smartphone takes the advantage of rapid increase of mobile users in emerging economies like India. In particular, recent advances in mobile computing and machine learning domain has led to research area in which Smartphone sensors can provide data for different types of applications e.g. road and traffic monitoring dynamically in spite of having static sensors on the road. The Smartphone is the best source for crowd sourcing real time dynamic updates while travelling. Smartphones are carried by a majority of individual these days, and are equipped with sensors like accelerometer, gyroscope, GPS, etc. effectively provide means to traffic sensing.

1.1 Research Motivation

The main motivation behind our research is the often extreme congestion conditions on Indian roads. This affects the commuters travelling incurring high travel time, frustrated mind-set, and high fuel consumption. The series of plots, see Figure 1.1, show that congestion varies with vehicle type. The average speed over the route and total travel time varies from vehicle to vehicle. As seen from the Figure 1.1(a), the 2-Wheelers (Bike) reaches the destination faster than car, followed by 3-wheeler, followed by public transport. The Bike (i.e. 2-Wheelers) has higher average speed and acceleration than other vehicles. In case of high congestion, the pattern varies significantly from vehicle to vehicle, while in case of smooth flow or low congestion, the total travel time for all the vehicles is almost same. They will reach the common destination at

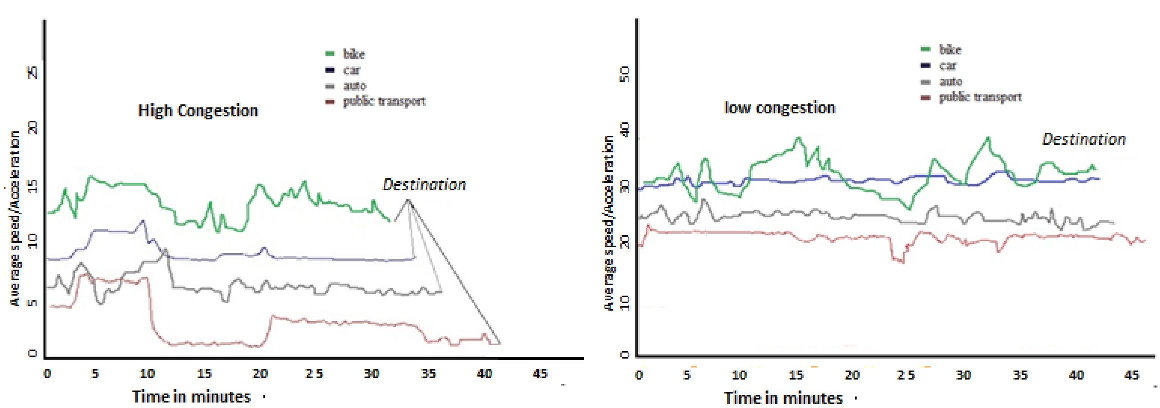

Figure 1.1: Travel time comparisons of different vehicles in high and low congestion

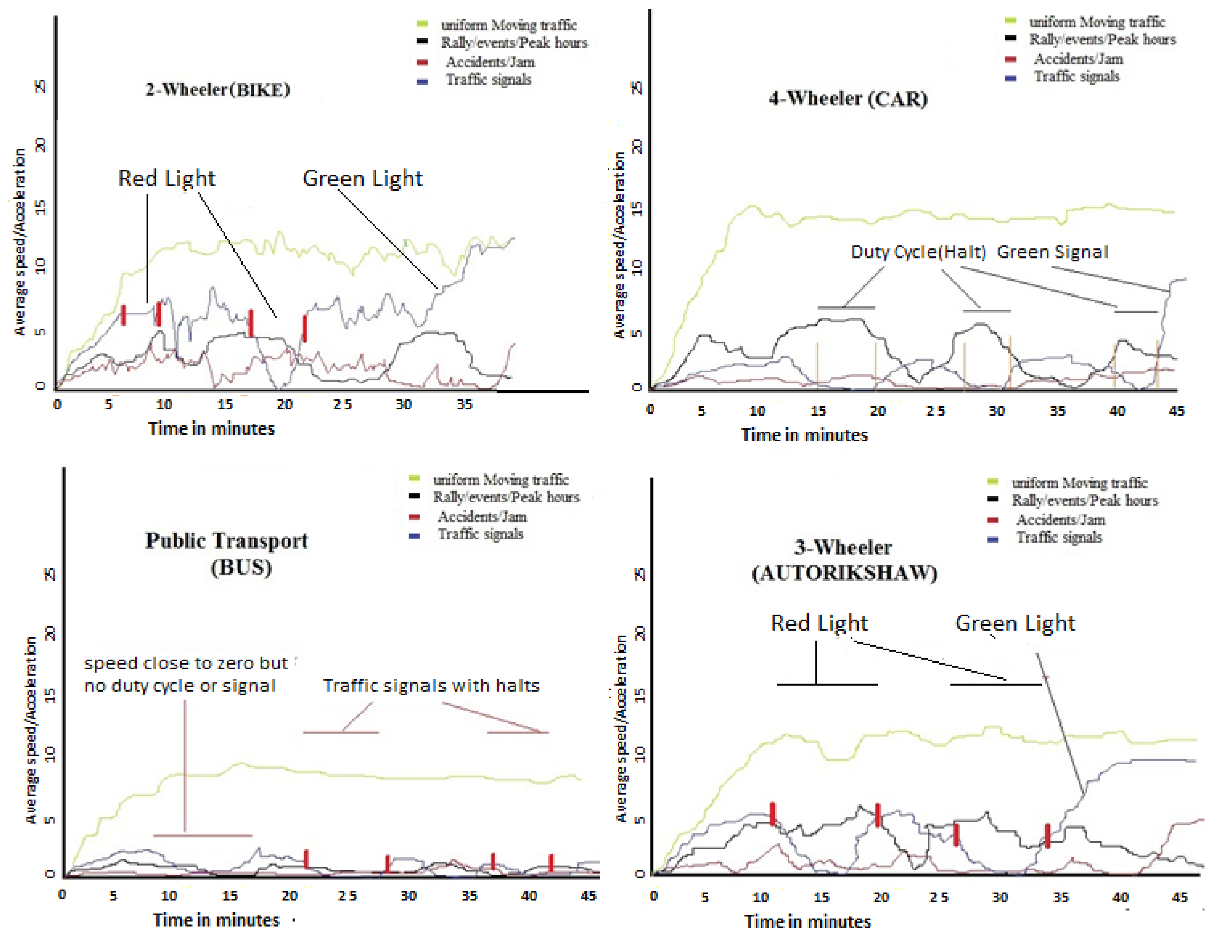

Figure 1.2: (a)-(d) clockwise from Top Left -Vehicle specific categorization for different traffic scenarios

the same time, though the average speed and acceleration varies to some extent. After seeing the following patterns from the data, it motivates us to study and research on factors for congestion based on different transportation modes.

The next series of plots, see Figure 1.2, give the motivation behind definition of congestion varies in different cases like complete blockage, moving traffic and traffic signal; hence need to be studied in different dimensions. All the traffic scenarios are plotted for different transportation modes against same feature (average speed/acceleration). As seen from the plots, in case of uniform moving traffic, the average speed is more smooth and constant. During rally (moving traffic), the average speed increases and decreases giving fluctuations and spikes in between. In case of accident (jammed), it fluctuates while suddenly the average speed increases in forward or backward direction. In case of traffic signals, it depends on duty cycle of traffic light. When green signal appear, the average speed increases suddenly. On looking at different patterns in the plots, it motivates our problem to study congestion based on different traffic scenarios. In our paper we focus the traffic sensing and monitoring based on different scenarios. On one hand, we give a novel approach for congestion detection; on the other hand we also make consideration for energy efficiency. Our proposed approach uses a cloud server for aggregation and classification.

It compares and classifies algorithm for vehicle classification and predicts traffic congestion level based on a particular traffic scenario. To make it more energy efficient, it uses GPS suppression and cloud offloading. The remaining paper is organised in sections as follows: Next Section provides the related work. After that we describes the Methodology employed, data collection and pre-processing. Next two sections describe the details about the vehicle classification and congestion detection algorithms respectively. Last section gives the discussion about the effect of road conditions on traffic congestion; energy efficiency, limitation and future goals. We end with the concluding remarks. The main contributions of this paper are: (1) a novel idea of research on congestion level relative to different vehicles using Smartphone sensors and present the intuition behind studying congestion of various types - blocked, moving traffic, near Traffic Lights, etc. differently. (2) Proposed an energy efficient algorithm to effectively find regions of traffic as per vehicle type. (3) Conducted experiments on the real world scenarios to verify the effectiveness of the algorithms.

1.2 Research Aim

The research aim of the work presented in this report is the following:

1. To increase our understanding of traffic and road anomalies and investigate effective solutions to combat the traffic congestion detection problem by mining smartphone sensor data.

2. To propose an energy efficient framework for the congestion detection level problem considering traffic, road and vehicle anomalies and other possible factors.

1.2.1 Advantages

The Work presented can be combined with other application or used as standalone application in case of congestion management to give following uses:

1. Suggesting optimal alternate route

2. Saving travel time

3. Reducing fuel consumption

4. Emission reduction

5. Frustration avoidance due to long queues.

Chapter 2

Related Work and Research Contributions

2.1 Related Work

There has been extensive research in the area of vehicle classification and traffic monitoring.The work presented in this report belongs to the area of congestion level detection on Road Traffic. In this section, we discuss some closely related work (to the experiment presented in this report) and present novel research contributions in context to existing work. We categorize the related work in 3 areas Vehicular Classification, road anomalies detection and Traffic anomalies detection.

2.1.1 Road anomalies detection

Chawla et al. [7] proposed a framework in which they modelled the road structures as a directed graph then employed PCA algorithm to detect road anomalies. Thajchayapong et al. [5] uses Gaussian process to model the microscopic traffic variables such as relative speed and inter-vehicle spacing which in turn were used to infer the traffic anomalies. The Pothole Patrol system [11] performs road surface monitoring by using special-purpose devices with 3-axis accelerometers and GPS sensors mounted on the dashboard of cars. Mohan et al. [18] focuses on novel aspects of sensing varied road and traffic conditions, such as bumpy roads and noisy traffic. Furthermore, they [18] uses Smartphone that users happen to carry with them, depending only on capabilities that are already available in some phones and that are likely to be available to many more people and geographical places in the coming years.

Begin of Table		
Study/Projects	Type	Goal/Objective
P. Mohan et al. [18]	Road and Traffic Anomalies detection	By using smartphone sensors and combining the GSM localization, GPS and accelerometer data it focus on monitoring road anomalies and traffic conditions. They check for the road quality and traffic localization, however do not provide any algorithm for route mapping or travel time estimation.
Eriksson et al. [11]	Road anomaly detection	It uses a 3-axis accelerometer and GPS enabled device mounted on the dashboard of the car to monitor different types of road anomalies like potholes, manholes, etc.
Thajchayapong et al. [5]	Traffic anomalies detection	They present algorithm that detects transient changes in traffic patterns using microscopic traffic variables such as relative speed and inter vehicle spacing. They use the Gaussian process to model traffic variables to detect traffic anomalies.
Chawla et al. [7]	Road Anomaly Detection	They proposed a framework to analyze a large GPS data set and employed PCA algorithm to detect anomalies
Pushpendra et al. [23]	Driving patterns	used in-built smartphone sensors like GPS, micro-phone and accelerometer, to detect driving behavior along with road and traffic conditions.

Continuation of Table 2.1		
Study/Projects	Type	Goal/Objective
Hemminki et al. [14]	Vehicle Classification	They used smartphone sensor-accelerometer features to classify the transportation modes. They gave novel features and classifiers which detect using smartphone data to detect and classify stationary, in motion and type of transportation modes. They also used adaptive boosting for combining the classifiers
Reddy et al. [21]	vehicle classification	combine GPS and accelerometer to recognize between stationary, walking, running, biking and motorised transportation, achieving high, over 90% accuracies. Classification is performed with a hybrid classiffier consisting of a decision tree and a first order discrete HMM classifier.
Zheng et al. [25]	vehicle classification	created and tested a number of classifiers, varying the segmentation method (fixed length, fixed duration and change point based), the generated features and the classification algorithm as well as introducing post-inference enhancement techniques using a database of change points combined with a DHMM. Classification in all cases is into classes walk, car, bus, biking
Manzoni et al. [16]	vehicle classification	Created a classification system that using only accelerometer data classifies into bus, metro, walk, bicycle, train, car, still, motorcycles with an accuracy of 82.15% percent.

Continuation of Table 2.1		
Study/Projects	Type	Goal/Objective
Jules et al. [24]	Traffic Monitoring	Presented a model to automatically detect traffic accidents using accelerometers and acoustic data, immediately notify a central emergency dispatch server after an accident, and provide situational awareness through photographs, GPS coordinates, VOIP communication channels, and accident data recording
UC Berkeley Millennium project [4]	Traffic Monitoring	Built software to report traffic delays on mobile phones and use mobile phones as probes for detecting traffic. Focuses on real time traffic monitoring
SmartTrek project [9]	Traffic Monitoring	Use vehicle based GPS client server model to track the movement of vehicles
OnStar project [3]	Vehicle Motion Detection	Tracks the movement of vehicles using GPS-enable units. It reports and analysis the information to a server for aggregation

Table 2.1: Related Work

2.1.2 Activity & Vehicle Classification and Detection

Pushpendra et al. [23] detect the driving behaviors using smartphone sensors. B.Pai et al. [20] Describes the problem of detecting and describing traffic anomalies like accident, celebrations, disasters etc. using crowd-sourcing using human mobility from driver routing behaviour and social media. Hemminki et al. [14], Reddy et al. [21], Zheng et al. [25], David piggott in his dissertation [8], and Manzoni et al. [16] use Smartphone sensors and classify vehicles.

2.1.3 Traffic Anomalies Detection

Due to immediate need and essence of the traffic monitoring and management approaches, research is actively going on in both commercial spaces Cartel project [15], SmarTtrek system [9], Bangalore traffic intelligent system [1], GM's OnStar [3]) as well as in academics. [3] Use vehicle based GPS client server model to track the movement of vehicles. While [15] use GPS based hardware installed in vehicles for route planning, SmartTrek [9], BTIS [1] use static sensors like loop vehicle detectors, traffic cameras, Doppler radar etc. for traffic prediction. Jules et al. in WrekWatch [24] uses cellular radios for the communication and monitor trains, buses etc. There has been work on Smartphone based accident detection systems. The Mobile Millennium project at UC Berkeley [4] has built software to report traffic delays on mobile phones, as well as to use mobile phones as data nodes for detecting traffic conditions on real time, providing almost real time reporting for the same.

2.2 Research Contribution

In context to the closely related works, this report makes the following research contribution:

1. The work presented in this report is the first step in the direction of detecting traffic anomalies based on vehicle types. Although there is lot of research in the vehicle classification and traffic anomalies detection, our work consider them in sequence and present a novel area of research.

2. We conduct empirical analysis on real world dataset acquired from smartphone sensors to train and test the effectiveness of the proposed vehicle type as features.

3. We also considered the effect of taking different traffic scenarios and coupled it with the vehicle types to conduct our experiments.

Chapter 3

Data collection and Proposed Solution Approach

3.1 Data Collection and Preprocessing

We collected real time crowd sourced Smartphone sensors data from the two locations in Delhi covering more than 1500 km. of both interior and outer areas. The sensors involved in data collection were accelerometer, linear acceleration, Gyroscope, Orientation, magnetometer, Light intensity meter, proximity, sound level and GPS. We use LG Optimus 4x, Google Nexus 4 and Samsung galaxy Ace phones having accelerometer accuracy as 1g each. The Smartphone X-axis represents the left-right, Y-axis as top-down, Z as front-back sides of phones. The areas covered were the IIT Delhi campus (interior) and Ring road area Punjabi Bagh - Lajpat Nagar (Via ring road). We covered both typical street areas and broader city roads with both controlled and uncontrolled traffic conditions. In IIT-Delhi campus, we collected sensors data using Car (Tata Indica), 2-Wheelers (Hero Honda CBZ extreme/ Honda Activa) and 3-Wheelers (3-wheeler), but we used all types of vehicles (2-Wheelers/Car/3-wheelers/Bus) on Ring Road. We collected data in IIT Delhi interior using 3 vehicle types for 10 days with 10+ km. each day. The same data was collected in the Ring road stretch by all 4 types for 10 days with 30+ km each day, see Figure 3.1.

During the data collection we noted the level of traffic as high, medium and low from the Google maps. At present, we took traffic scenarios in the ground truth. We foresee the detection and classification of traffic scenarios as our next goal. The choice of data collection roads was motivated by rich traffic conditions like traffic lights, flyovers, accident prone areas, business routes, etc. Each of the two types of data collection was done in the daylight. For appropriate analysis, all the vehicles followed the same route from the same source.

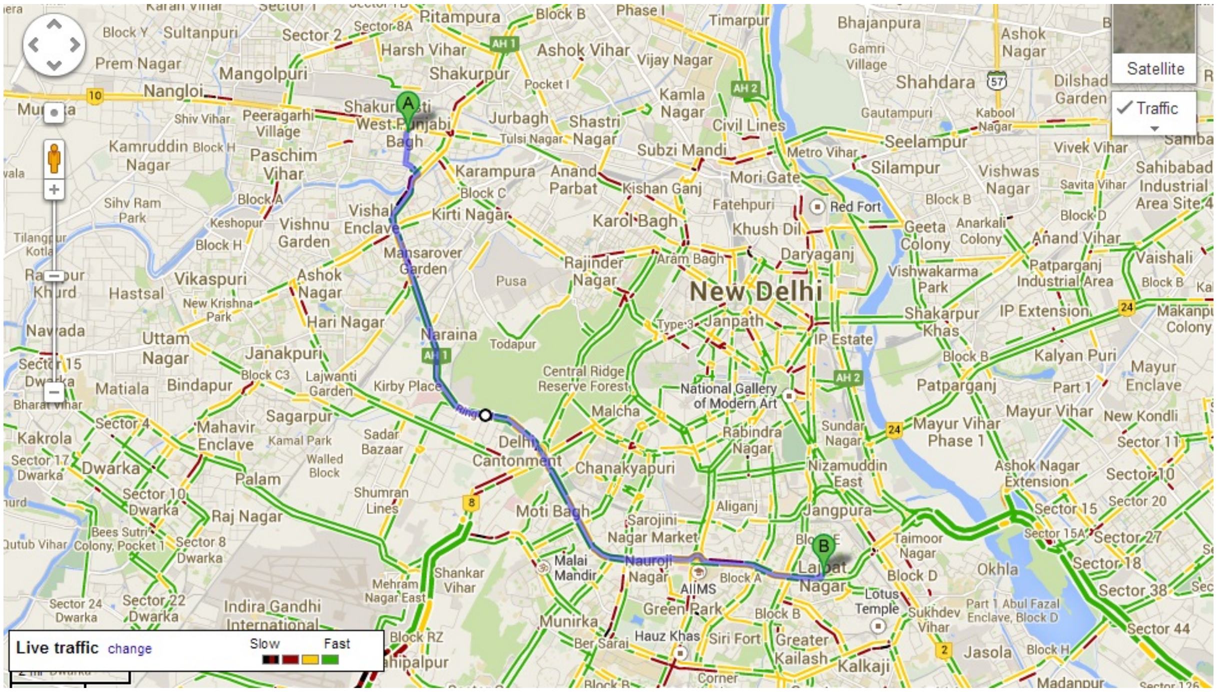

Figure 3.1: Data Collection on Ring Road

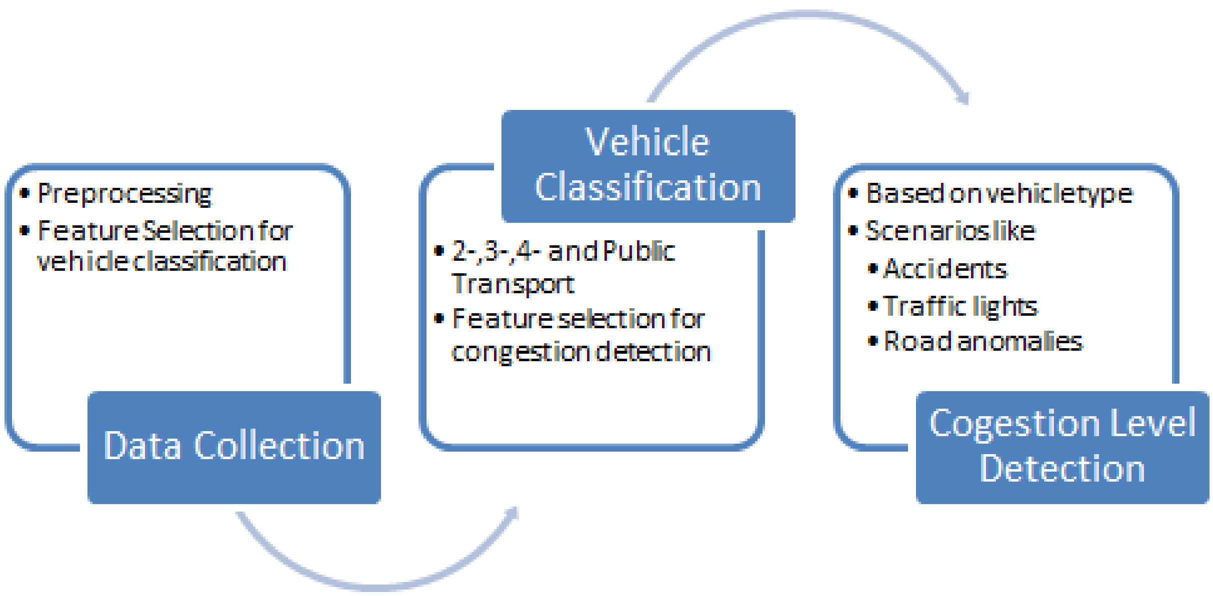

Figure 3.2: Block Diagram

3.2 Solution Approach

Figure 3.2, gives the complete work flow of the algorithm. First data collection is done using the Smartphone sensors. Next step is feature selection which is input to the classifiers used for transportation mode classification. Finally, along with features extracted and vehicle type, categorized as per traffic scenarios; traffic congestion level detection is done. After the feature extraction (see subsequent sections), the data is passed through low pass and high pass filters to distinguish the gravity component. The filters consider the sudden changes during motion. After passing data through filters, it is checked if vehicle is stationary or moving. For this, we did statistical analysis using 3-axis acceleration. Below is the algorithm 1 for vehicle motion detection:

Input:Re-sampled 3-axis Accelerometer Sensor data, Threshold, T
Result: Vehicle Motion Detection(True = moving)
initialization : windows per sample=N, i=10,11.. 20;
Window size: $\Delta_i = [10, 20]$
while $\Delta_i \leq 20$**do**
  total acceleration, $T_{acc} = \sqrt{(x^2 + y^2 + z^2)}$
  for each window compute range, R $= max(T_{acc}) - min(T_{acc})$
  if *R persists over the consecutive windows and R*$\leq T$ **then**
    return True;
  else
    return false
  end
  $\Delta + +;$
end

Algorithm 1: Vehicular Motion Detection for varying window sizes

Once the vehicular motion is confirmed the vehicle classification is performed followed by the congestion detection. We detect the effective features for vehicle type and congestion level detection. It considers time domain analysis and frequency domain analysis. Evidence obtained by exploratory study from the literature enabled us to make a corpus of features. It includes time domain statistical features like mean, median, correlation, cross-correlation, min-max, variance, std. deviation, etc., frequency domain features like FFT, Energy, dominant frequency etc. and discrete domain analysis like dynamic time wrapping. The data is segmented in different window sizes for all the analysis. The main part of the detection algorithm is choosing the right window size for feature extraction and classification. The window size is chosen so that it can detect the sudden changes/tweaks. Using above features, we classify the vehicle type using various known classification algorithms. We also compare and analyse the classification accuracy. We use ensemble based learning algorithm like bagging and boosting for traffic classification.

Chapter 4

Empirical Analysis and Performance Evaluation

4.1 Vehicle Detection and Classification

4.1.1 Feature selection and importance

The accelerometer and GPS, in combination, is widely used and compared [21]. However, GPS has many limitations like high power consumption, GPS receiver's dependency on obstructed view to satellites and unreliable and low accurate solutions for transportation mode behaviour [14], [19]. For feature importance we used two methods and cross- validated their observations. First, we took all features and applied uniform removal approach to find the important features. Second, we carefully choose the features using correlation based feature selection method (CFS) [13] considering both the features taken from the literature [14], [21] and new features which are not explored in the literature. We carry on the footprints of [21] for vehicle classification accuracy and applied the well-known classifiers. We, however differ from the mentioned approach in different set of features tested and different class of vehicles. Table 4.1 summarizes the features and sensors with literature references.

4.1.2 Classification

We compared and analysed different classifiers like C4.5 decision tree classifier(J48), Naive-Bayes classifier (IBK), K-Nearest neighbour classifier, HMM classifier (Custom source and embedded in Weka), Support vector Machine (SVM). We take 2150 window instances of vehicles out of which 678 2-Wheelers, 443 cars, 715 3-Wheelers, 314 Public Transport instances were recorded as per different time. The confusion matrix for the same is depicted in the table 4.3.

Sensors Specifications	Description and Features	Reference
Accelerometer	Mean, median, min, max, linear Speed, variance, energy, FFT coefficients	[14], [21], [25], [8] [16]
Magnetometer	Magnetic field	[8]
GPS	Longitude, Latitude, linear Speed	[19], [21], [25], [8]
Gyroscope	Rotational speed	[8]
Light Intensity	Light falling on the device	[8]
Microphone	Sound from the surroundings	[8]
Temperature	Temperature from the surroundings	[8]
Orientation	Orientation of the phone	[8]

Table 4.1: Feature Survey

4.1.3 Results and evaluation

Comparing Classifiers of vehicle classification: We compared different classifiers for vehicle classification mentioned in [21] for our data set. The training and testing of classifiers was done using a 3-fold validation in case of IIT collected data and 4-fold validation in case of Ring road data. We have evaluated above mentioned classifiers in classification algorithm and execute in Weka version 3.7.9 [12]. The choice of classifiers was done based on exploratory analysis with related work and type of features used in our approach. The following evaluation matrices were used for evaluation of the classifiers performance: Precision or positive predicted value (PPV), Recall or sensitivity, negative predicted value, specificity and accuracy.

From Figure 4.1, it can be seen that acceleration and orientation are best features as expected. The classification accuracy does not get affected much with light, temperature. We perform various experiments and concluded that temperature, light, sound are not very good features. The acceleration, gyroscope are also good features. The most important features are acceleration and orientation. Classification accuracy as shown in the table 4.2 highlights the C4.5 decision tree as the suitable classifier with more than 90% vehicle wise accuracy and 85% accuracy overall.

Classifier/ Vehicle type	2-Wheelers	Car	3-Wheelers	Public Trans-port	Total
Naïve Bayes	70.86	71.34	75.67	73.24	71.67
C4.5 (DT)	92.69	93.76	90.41	94.65	85.76
K-NN	73.87	79.23	77.45	75.23	76.45
HMM	82.67	81.22	79.25	78.98	80.53
SVM	76.50	75.23	74.32	78.23	76.32

Table 4.2: Classification Results for different classifiers

Actual/Predicted	2-Wheelers	Car	3-Wheelers	Public Trans-port
2-Wheelers	603	20	43	12
Car	36	376	26	5
3-wheelers	43	23	600	49
Public Transport	3	24	22	265

Table 4.3: Confusion Matrix for vehicle classification

Figure 4.1: Classification accuracy with different sensors combinations

4.2 Traffic Congestion Detection and Classification

4.2.1 Feature selection and importance of Various Traffic Scenarios

There are different traffic anomalies like accident, rallies, traffic signal and road anomalies etc. All these sources create two major variants of traffic as moving traffic or complete blockage (Low congestion vs. high congestion). Both of these situations have different characteristics of traffic. The average speed in moving traffic is different than in blocked traffic. The travel time to reach a destination is different in both. For feature extraction, the algorithm considers the sensors data of accelerometer, GPS, gyroscope and Magnetometer with statistics like mean, variance, energy, FFT, etc. of acceleration and speed in spatio-temporal context.

Uniform Moving Traffic

In this case, the 2-Wheelers and car almost follow the same pattern. The 2-Wheelers may show little fluctuations in case of high congestion but the motion is almost smooth, see Figure 4.2.

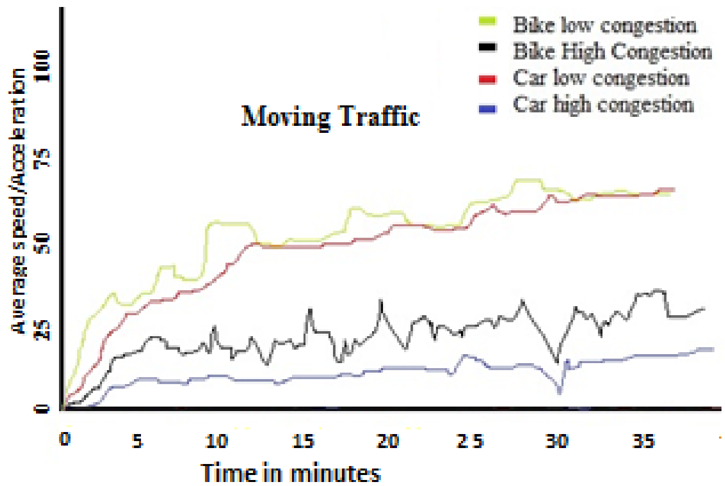

Figure 4.2: Average speed/acceleration variation for different level of traffic congestion in case of car and 2-Wheelers.

Traffic Signals

Our detection algorithm for different types of vehicles for high traffic works as follows:

1. *Input:* duty cycle of traffic signal (5 minutes in our case), we fixed different window sizes as per a time window e.g. 10 min, 20 min, etc.
2. The relative speed between consecutive windows is greater than certain threshold. This threshold is chosen based on the average of speed per window per vehicle type.
3. This is facilitated by the fact that the difference between location values of vehicle and traffic signal does not move very further from zero within 1 - 2 duty cycles in the case of highly congested traffic.
4. This is also affected by the position of vehicle and distance from the traffic lights.
5. In all such state of affairs, it can be one of the hot-spot of traffic.

Algorithm 2: Vehicular Motion Detection for varying window sizes in traffic lights

Figure 4.3 shows the low congested and highly congested traffic for 2-Wheelers and car.

The next plots, see Figure 4.4(a) and (b), show the occurrence of duty cycle when the vehicle is near or far from the traffic signal.

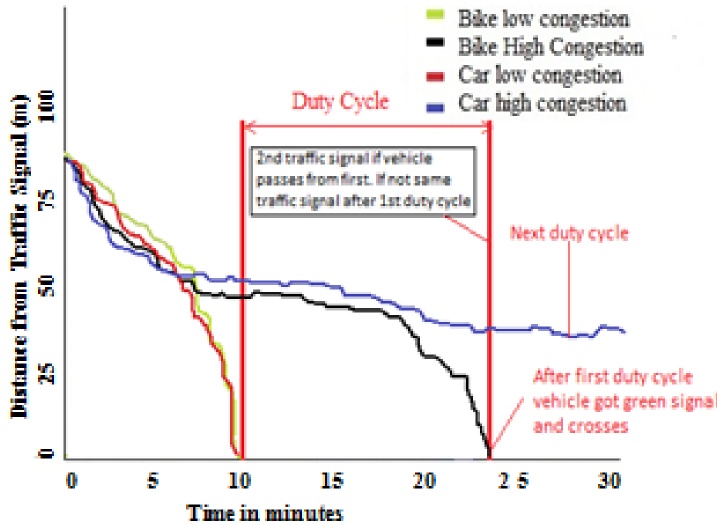

Figure 4.3: Traffic signal with duty cycles in case of 2-Wheelers and car.

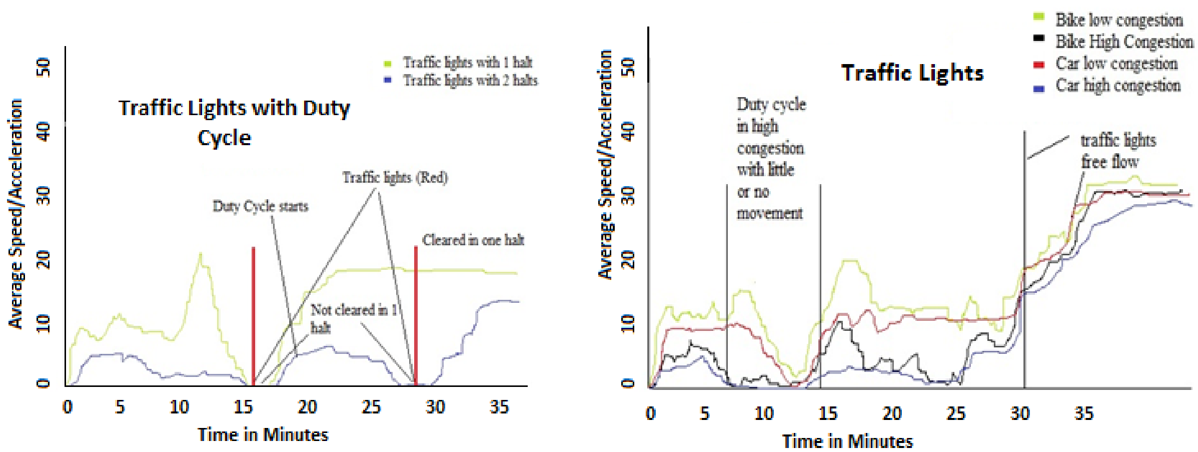

Figure 4.4: vehicle localization near or far from traffic signal and duty cycle occurrence.

Business peak hours, Election Rallies and celebration events like marriage

We analyse these scenarios using acceleration, speed, and compass Smartphone sensors. These sudden/occasional events are periodical and planned generally in terms of time but are susceptible to randomness caused with unaware location and thus create unpleasant low-moving traffic for commuters of the same path. We performed various experiments using different features to analyse the effects of this scenario. Due to the randomness and unplanned behaviour, we repeated our experiments by taking the 50% overlapping window size for the data per vehicle type. We select the threshold on the vehicle speeds as follows:

1. Calculate min, max and average speed of each vehicle for each window.

2. Compute the similarity measure using spearman's rank coefficient matrix for each range and vehicle type.

We observe that relative speed between consecutive windows that crosses threshold value for 2-wheelers during stop and go behaviour is more than 4-wheelers. Also, rate of change of compass reading is more in case of 2-, 3- wheelers as compared to 4- or more wheels vehicles. An important characteristic observed in the case of compass sensor is that 2-Wheelers will exhibit more fluctuations in case of high congestion while other vehicles have occasional fluctuations in case of low congestion only and be uniformly constant in high congestion traffic. This is because to reduce travel time 2-wheelers weave through traffic and make more movements than in low congested traffic. However, even though relatively heavy vehicles also strive to weave through traffic they are often stuck due to their large size. This suggests that accelerometer and compass/phone orientation can be useful in these kinds of scenarios. As expected, we observe this pattern arising for more than 20-30 min is leading with high chance to it to be predicted as one of the region affected.

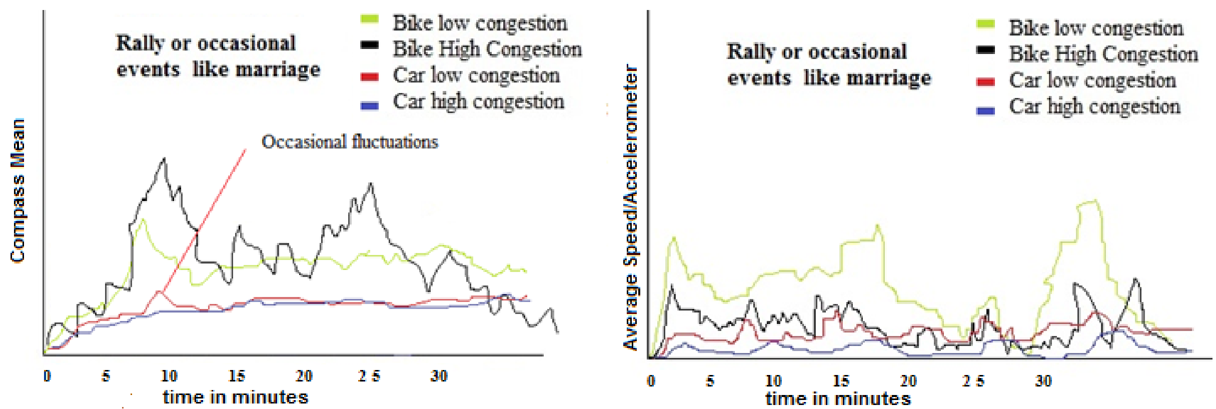

Figure 4.5: Using Compass and accelerometer as feature fluctuation in the case of rallies, celebration events and peak hours under different congestion levels for 2-Wheelers and Car.

Figure 4.5(a) and (b) shows the smooth flow and high traffic features during business peak hours and rallies and corresponding congestion level as per vehicle type for accelerometer and compass mean.

Accidents and Jams

In this case there is utter blockage of roads. In India, there are side-way paths for pedestrian and 2 wheelers to pass by. Moreover, there are some roads of shorter width from which only 2 wheelers can pass by. The pattern of stop and go with peaks and fluctuations varies for different transportation modes. The 2 wheeler will pass by, whereas the car will become stationary and wait. This pattern can be seen in the Figure 4.6 using statistical features on acceleration, compass, and speed mobile sensors. With reference to [17], we analyse the deceleration behaviours. In many cases heavy vehicles like 4-or more wheelers, if they are free from backward traffic, they try to reverse back for finding new routes.

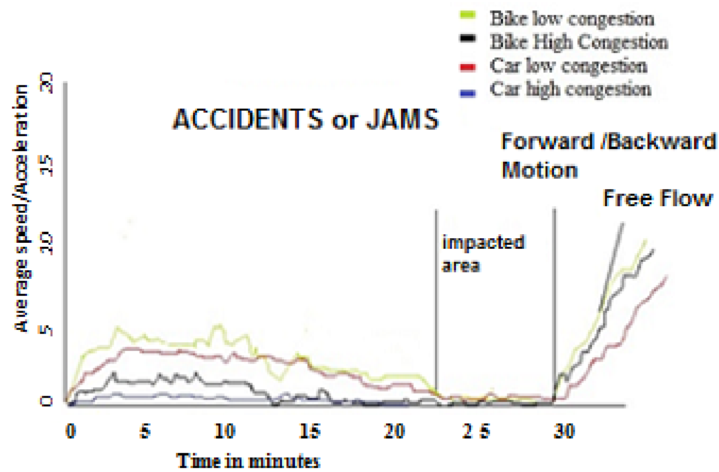

Figure 4.6: Forward or backward acceleration in case of Accidents and Jams for Car and 2-Wheelers in different level of congestion.

4.2.2 Classification

After appropriate extraction of features of all the above cases, we chose ensemble based classification approach. In spite of doing single decision using algorithms like decision tree, it is better to take decision from ensemble of classifiers for this problem. The data is re-sampled again and again with different window sizes using bagging algorithm. The final decision is taken based on average of the decisions. We use AdaBoost, which is ensemble learning algorithm that learns from the weighted examples. Initially, each sample is given weights uniformly. The error is calculated and weights are updated for each sample as per error. It is assumed that the error should be less than 0.5. The final classifier is based on the weighted majority voting. To quantify the performance of this algorithm, the magnitude of margin is a good measure. The margin is defined as the difference between weighted fraction of weak classifiers predicting the correct label and the weighted fraction predicting the incorrect labels. Higher the margin, higher is the confidence of the algorithm. Window size was selected based on vehicle type in each scenario type; decision tree is made based on information gain, root being the most important feature. All the respective features like accelerometer, GPS, compass, speed, etc. are given weights, and the most important feature is given the highest weight in each traffic scenario respectively. We normalize the weights to common scale for all the traffic scenarios. Based on these feature weights, congestion level classification is done using decision tree. Then the miss-classification is calculated. The weight of the miss-classified points is increased. This complete algorithm is repeated for 10 times with majority weight adjusting. The final classification is done using ensemble of all the weak classifiers (decision trees). This is an improved algorithm over already existing AdaBoost algorithm. We also made comparison with single decision tree with our new classification algorithm; the classification accuracy improves by 5%.

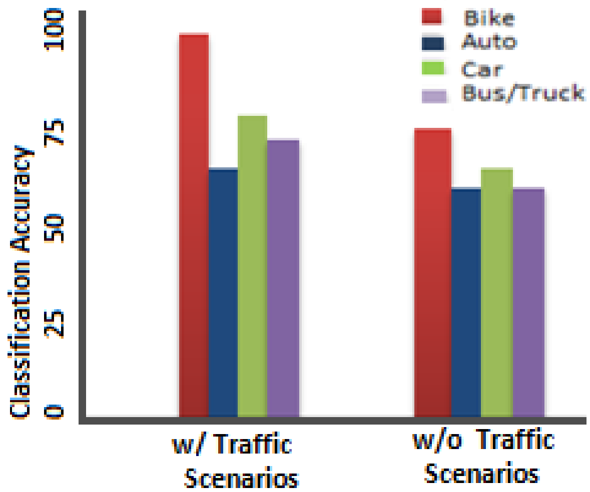

Figure 4.7: Effect of vehicle types on classification accuracy.

4.2.3 Results and evaluation

For congestion detection we employ vehicle based window strategy. After vehicle classification is done we recognize the pattern of vehicles on the particular route and then segment them as per their type. We segment the vehicle prediction of a particular type constitutes 3/4Th of the window size. This is done to compensate for inaccurate vehicle classification. For example if vehicle type classification is two-wheelers, 2-Wheelers, car, 2-Wheelers, 2-Wheelers, car, 2-Wheelers, 2-Wheelers, and 2-Wheelers; this window is annotated as a "2-wheeler" or "Bike window". Once the window is annotated, we select the window for classifying traffic congestion level. We detect the level of congestion as per vehicle type. For evaluation, we compare the predicted traffic congestion level to the ground truth. Figure 4.7 shows the high congestion classification accuracy considering and not considering each vehicle type. This plot shows the importance of considering vehicle type while studying high congestion. Figure 4.8 shows the classification accuracy by not considering traffic scenarios and taking combined features. Figure 4.9 shows the high and low congestion classification accuracy using both vehicle type and traffic scenarios.

As shown in Figure 4.7 and Figure 4.8 give classification accuracy for high congestion. Considering vehicle type and traffic scenarios, the congestion level (high and low) classification accuracy is evaluated. We include both vehicle type and traffic scenarios in our training dataset and compare it with test dataset neglecting both the features. We were able to predict congestion level with more than 80% accuracy. We use evaluation metrics as: precision, recall, false positive rate and true negative rate. We have considered sensors like accelerometer, orientation magnetometer, GPS which we argued to predict good classification attributes. The plots show the recall for each feature. As seen from the Figure 4.10,

the features like average total acceleration, standard deviation of acceleration came out to be a good feature. From Figure 4.11, the compass mean and GPS difference between nearest traffic

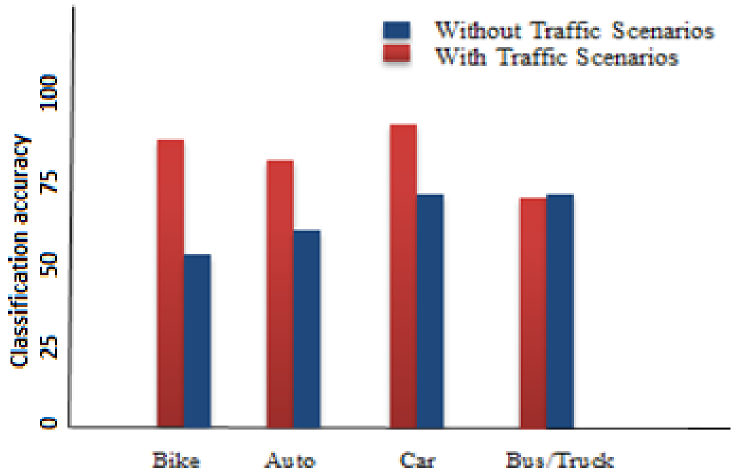

Figure 4.8: Effect of vehicle types on classification accuracy.

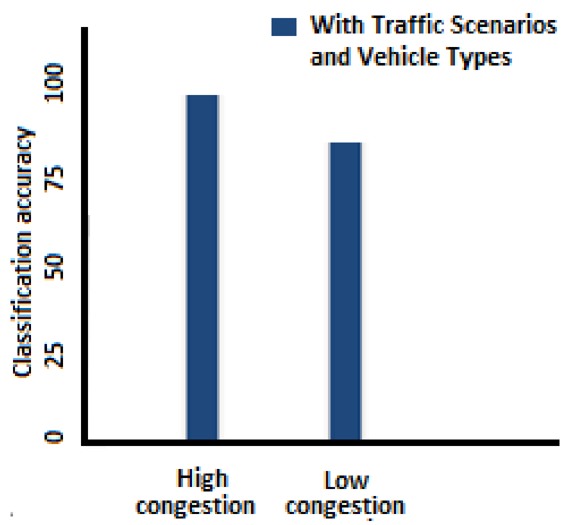

Figure 4.9: Classification accuracy of congestion level taking both vehicle types and traffic scenarios.

signal and vehicle position came out to be a good features. These features came out as expected. We implemented the complete system in real time scenarios in Delhi. We found that vehicle classification latency is not more than 5-10 minutes. We integrated this with Google Maps and can able to tell traffic as per vehicle type at different times of day. We can make more accurate results of congestion levels per vehicle type by considering different traffic scenarios. The empirical results show the scenario where at the same time and route, the congestion is for car, truck, but not for 2-Wheelers.

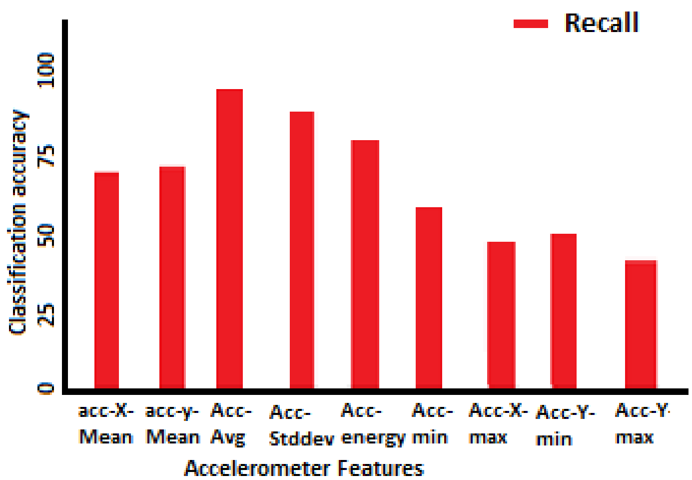

Figure 4.10: accelerometer sensor feature importance and classification accuracy.

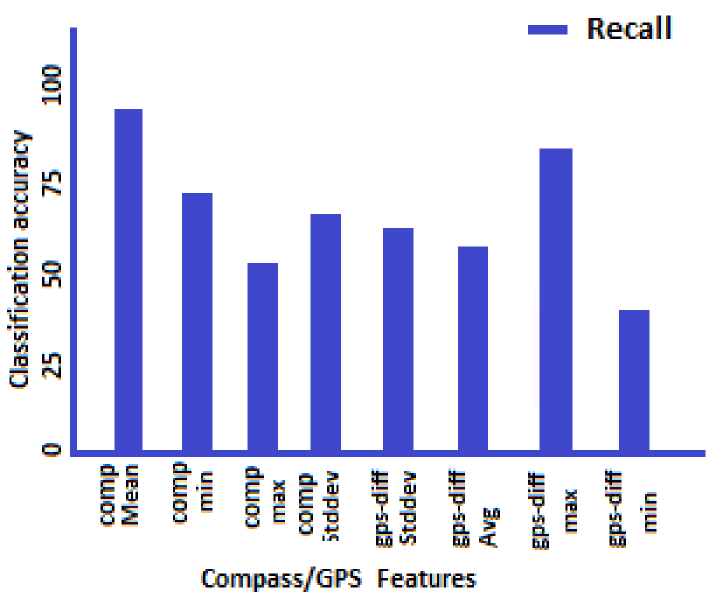

Figure 4.11: Compass and GPS sensor features importance and classification accuracy.

Chapter 5

Discussion and Conclusion

5.1 Effect of Road anomalies on Traffic Congestion

There are various road anomalies like pothole, speed breaker, turns, intersections; rough road; which can lead to decrease in decrease in average Speed of the vehicles. The 2-Wheelers will bump more and have high centripetal acceleration than any other vehicles. There is a change in congestion level in both smooth road and different road anomalies. Different vehicles behave different in different road anomalies. This directly influences the congestion level, since higher is the road anomaly higher is the chance of congestion than normal. In case of rough road, even the car will try to make turns to get fewer potholes on the way, the average acceleration varies considerably, and the bus/public transport will suffer less roughness of the road. The average acceleration of the bus does not fluctuate as much. 3-wheelers being unstable, their average acceleration value fluctuates the most on a rough road. Figure 5.1 show the valley-spikes in case of road anomalies and decrease in acceleration.

Figure 5.1: Road anomalies and Decrease in Acceleration.

5.2 Energy Efficient Framework

In case of vehicle classification using road conditions, we are using four different principles for improving energy consumption- Substitution, Suppression, Piggybacking and Adaptation [26]. Substitution means including other low consuming sensors for location estimation. Suppression uses sensors like accelerometer. Piggybacking synchronizes the location updates requests from multiple running location based applications. In case of road anomaly effect on congestion level:

- Vehicle moves on smooth road
  - GPS is switched off.
- Occurrence of Pothole, breaker or turn detection by doing accelerometer analysis.
  - the GPS is turned on till the end of rough road.
- This pothole or any road anomaly is approximately for 1 meter.
  - leads to consumption of energy.
  - Solution:The vehicle classification is achieved using less GPS, and deriving speed from accelerometer which is not energy consuming, with Other sensors like light, sound, temperature sensors which are less energy consuming.
- For classifying traffic congestion
  - we rely on cloud offloading. During the same state e.g. moving or jammed, location estimates are done using offline cloud methods from acceleration except near traffic signal part.
  - In a sense, the same traffic state location estimates are done using offline cloud computing whereas near traffic signal congestion state location estimates is done using GPS.
  - GPS has to track for few seconds only.

5.3 Limitations and Future Work

Although we categorize the congestion level on basis of vehicle type, there are some conditions which we have either not addressed or assumed to be satisfactory:

- We have assumed that the Smartphone is kept in the front pocket of the trouser parallel to the ground for the entire time. There are known approaches in literature to account for Smartphone orientation. We have not incorporated them in our data collection and analysis.

- We have not taken the scenarios where the Smartphone is carried by a pedestrian. Hence, it avoids the case where the pedestrian prediction can be confused by highly congested traffic.
- The latency of our approach is limited to minimum 5 minutes of travel. In the scenarios where the travel time is less than 5 minutes, our system is susceptible to no recognition of congestion levels and hence defeats the purpose.

Apart from overcoming the limitation mentioned above we have the following future goals:

- New features driven algorithm for vehicle classification considering the odd cases like change of transportation mode from vehicle to pedestrian and vice-verse.
- Propose and Implementation of more intense energy efficient framework along with the road anomalies considerations.
- Propose a generalized hot-spot detection algorithm considering all traffic scenarios comprising rural and urban locations.

5.4 Concluding Remarks

In this thesis, we have provided experimental evidence to demonstrate the feasibility of congestion as per vehicle type using vehicle classification. With the experimental insights, we develop a practical framework that extracts vehicle specific features as signatures to suggest congested areas specific to vehicle. In particular, we explore the time and frequency to select the most distinguishing features to feed into pattern matching algorithm. Then, adaptive algorithm is applied which can classify the features well. We advocate the use of vehicle type and traffic scenarios as a feature for traffic congestion. Our experimental evaluation shows that this approach can predict areas of congestion as per vehicle type with more than 80% accuracy.

Bibliography

[1] Bangalore Transport Information System. http://www.btis.in/.

[2] Intelligent Transportation Systems. http://its.dot.gov/.

[3] OnStar by GM. http://onstar.com/.

[4] The mobile millennium project. http://traffic.berkeley.edu/.

[5] Barria, J., and Thajchayapong, S. Detection and classification of traffic anomalies using microscopic traffic variables. *Intelligent Transportation Systems, IEEE Transactions on 12*, 3 (Sept 2011), 695–704.

[6] Bhoraskar, R., Vankadhara, N., Raman, B., and Kulkarni, P. Wolverine: Traffic and road condition estimation using smartphone sensors. In *COMSNETS* (2012), K. K. Ramakrishnan, R. Shorey, and D. F. Towsley, Eds., IEEE, pp. 1–6.

[7] Chawla, S., Zheng, Y., and Hu, J. Inferring the root cause in road traffic anomalies. In *ICDM* (2012), pp. 141–150.

[8] D. Piggott. Inferring Transportation Mode using Smartphone Sensor Data. Dissertation submitted for Computer Science Tripos, Fitzwilliam College. https://www.piggott.me.uk/wp-content/uploads/2011/08/Inferring-Transportation-Mode-using-Smartphone-Sensor-Data.pdf.

[9] Dailey, D. Smart Trek: A Model Deployment Initiative. WSDOT, 2011. http://www.its.washington.edu/pubs/smart_trek_report.pdf.

[10] Dash, D. K. India loses Rs 60,000 Crore due to traffic congestion: Study. http://timesofindia.indiatimes.com/india/India-loses-Rs-60000-crore-due-to-traffic-congestion-Study/articleshow/13678560.cms.

[11] Eriksson, J., Girod, L., Hull, B., Newton, R., Madden, S., and Balakrishnan, H. The pothole patrol: Using a mobile sensor network for road surface monitoring. In *Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services* (New York, NY, USA, 2008), MobiSys '08, ACM, pp. 29–39.

[12] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. The weka data mining software: An update. *SIGKDD Explor. Newsl. 11*, 1 (Nov. 2009), 10–18.

[13] Hall, M. A. Correlation-based feature selection for discrete and numeric class machine learning. In *Proceedings of the Seventeenth International Conference on Machine Learning* (San Francisco, CA, USA, 2000), ICML '00, Morgan Kaufmann Publishers Inc., pp. 359–366.

[14] Hemminki, S., Nurmi, P., and Tarkoma, S. Accelerometer-based transportation mode detection on smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems* (New York, NY, USA, 2013), SenSys '13, ACM, pp. 13:1–13:14.

[15] Hull, B., Bychkovsky, V., Zhang, Y., Chen, K., Goraczko, M., Miu, A. K., Shih, E., Balakrishnan, H., and Madden, S. CarTel: A Distributed Mobile Sensor Computing System. In *4th ACM SenSys* (Boulder, CO, November 2006).

[16] Manzoni, V., Maniloff, D., Kloeckl, K., and Ratti, C. Transportation mode identification and real-time co2 emission estimation using smartphones. *SENSEable City Lab, Massachusetts Institute of Technology, nd* (2010).

[17] Maurya, A. K., and Bokare, P. S. Study of deceleration behaviour of different vehicle types. *International Journal for Traffic & Transport Engineering 2*, 3 (2012).

[18] Mohan, P., Padmanabhan, V. N., and Ramjee, R. Nericell: Rich monitoring of road and traffic conditions using mobile smartphones. In *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems* (New York, NY, USA, 2008), SenSys '08, ACM, pp. 323–336.

[19] Mohan, P., Padmanabhan, V. N., and Ramjee, R. Trafficsense: Rich monitoring of road and traffic conditions using mobile smartphones. In *Technical Report* (2008), Microsoft Research, MSR.

[20] Pan, B., Zheng, Y., Wilkie, D., and Shahabi, C. Crowd sensing of traffic anomalies based on human mobility and social media. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (New York, NY, USA, 2013), SIGSPATIAL'13, ACM, pp. 344–353.

[21] Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., and Srivastava, M. Using mobile phones to determine transportation modes. *ACM Trans. Sen. Netw. 6*, 2 (Mar. 2010), 13:1–13:27.

[22] Sen, R., Cross, A., Vashistha, A., Padmanabhan, V. N., Cutrell, E., and Thies, W. Accurate speed and density measurement for road traffic in india. In *Proceedings of the 3rd ACM Symposium on Computing for Development* (New York, NY, USA, 2013), ACM DEV '13, ACM, pp. 14:1–14:10.

[23] Singh, P., Juneja, N., and Kapoor, S. Using mobile phone sensors to detect driving behavior. In *Proceedings of the 3rd ACM Symposium on Computing for Development* (New York, NY, USA, 2013), ACM DEV '13, ACM, pp. 53:1–53:2.

[24] White, J., Thompson, C., Turner, H., Dougherty, B., and Schmidt, D. C. Wreckwatch: Automatic traffic accident detection and notification with smartphones. *Mob. Netw. Appl. 16*, 3 (June 2011), 285–303.

[25] Zheng, Y., Li, Q., Chen, Y., Xie, X., and Ma, W.-Y. Understanding mobility based on gps data. In *Proceedings of the 10th International Conference on Ubiquitous Computing* (New York, NY, USA, 2008), UbiComp '08, ACM, pp. 312–321.

[26] Zhuang, Z., Kim, K.-H., and Singh, J. P. Improving energy efficiency of location sensing on smartphones. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services* (New York, NY, USA, 2010), MobiSys '10, ACM, pp. 315–330.