

# Behavior of Vehicle based Traffic Detection System

SweSweAung, Thinn Thu Naing  
University of Computer Studies, Yangon, Myanmar  
swesweaung.ucsy@gmail.com, thinnthu@gmail.com

## Abstract

*As Traffic congestion is becoming an everyday facing problem in urban region, monitoring road and Traffic prediction system are playing an important role in the city life. The previous Traffic Prediction Systems were implemented depending on the road network sensor. These technologies had been prompted by the need of addressing to solve the problem installation and maintenance cost. Fortunately the dramatic technology innovation is carrying many crucial solution for Transportation agency to provide the relative services efficiently. This paper mainly emphasizes detecting Traffic condition by analyzing the behavior of vehicle primarily based on applying GPS enable Mobile phone and integrating the underlying Transportation network information and history data. The system is built into two parts: Client (Mobile device) and Cloud Backend Server. On the Client side, the system distinguishes whether the Mobile device carrier is taking a vehicle or walking. The Average Moving Filtering method and the measurement of total distance are utilized in analyzing mode of Transportation. The distance of two points (latitude and longitude) is computed by using Haversine Formula. On the Server side, it detects the Traffic status based on checking the behavior of vehicle based on the Client result by applying Bayes Classifier.*

## 1. Introduction

The population is continuously growing in big cities and at the same time this increasing amount is also catching up by the rising vehicle. Thus, these two developing groups consequently make the existing transport infrastructures create severe transportation problems almost every day. Likewise, the existing one is not able to handle the amount of large vehicles causing heavy congestion nearly daily. The Traffic congestion can be mainly due to the unpredictable happening on roads, such as car accidents, road repairing and extremely bad weather conditions. The poor road network infrastructure in developing countries may occur bag congestion. The recurring congestion can also be arisen occasionally during rush hour time. The constantly travel delay on road

consumes a large amount of fuel, increasing monetary losses and effecting on the city's environmental pollution, also make staffs and schoolchildren to be time wasting on a bus. Waiting too long time at junction or in the middle of Traffic where people are stuck for two or three hours every day may impact awfully on health of people being more and more irritated day by day. For this motivation, the development of Intelligent Transportation System (ITS) has become a vigorous research field and been attractive to many researchers for ages.

Intelligent Transportation System is a significant one to its relating problem area such as Traffic congestion as well as taking a vital role in the economy and its environment. The advance of innovated technology in communications, computing and electronic sensors is bringing an energetic dosage to Transportation community to solve its concerning problems and to serve new services to citizens. The technologies which the preceding researchers were mainly utilized are the network of fixed sensors, such as roadside sensors, Camera, HD traffic (traffic information service), and Traffic signals for dealing with the Traffic congestion problem. These tools need to have large amount of investment for maintenance and installation cost throughout the road network. The renovated technology that comes to replace the old technology and which could be used to cope with the needed massive investment in solving Traffic congestion is GPS\_enable Mobile device. The dramatic development of Mobile device in Mobile market cannot be predictable and about over two third of population have Mobile device is not amazement in these days. Today's Mobile phone reached in the Mobile market has got a rich set of functions and embedded robust sensor platforms, such as GPS, camera, microphone, accelerometer, digital compass, gyroscope, near field communications (NFC), WiFi, Bluetooth, and cellular capabilities. Luckily, these sensors are an essential one for implementing new applications across a widespread range of domains, such as safety, environmental monitoring, social networks, transportation, and healthcare. The benefit of Smartphones is to be able to define its current position, and then transfer it to the backend server through the

remaining cellular network. Besides, the increase development of these devices can provide the required Traffic clues at very low deployment costs.

Myanmar is one of the developing countries as well as becoming to face the continuous increment of population and vehicle. This problem also gives rise to traffic congestion in down town area in Yangon. However, there is no enough resources throughout the road for traffic system to enable prediction of current traffic jams. The main resources are road side camera, road sensor and HD traffic (traffic information service). However, these tools is very expensive in installation and keeping maintenance. One brilliant technique to be used instead is GPS enable mobile phone. In 2014, Myanmar has over 50 million population and about 85% of it are becoming to own Mobile phone. In 2004 and before that, one of ten persons of Myanmar's 50 million population had held Mobile phone. According to the point, the approach of tracking the condition of Traffic by employing GPS\_enable Mobile device becomes tremendous method in monetary saving and reduction in human force utilization. However, the several challenges remain to be revealed in the GPS\_enabled Mobile phone based Traffic system. A large variety of potential applications for such systems are very rely on counting the number of active vehicles on the road for the prediction of Traffic situation. In other words, almost Traffic state is predicted by processing over the all available GPS logs. Typically, these likewise systems very depend on enough GPS clue for computing congestion appropriately. If only one vehicle can submit it's data to the estimator, although there is a Traffic jam, the system couldn't detect the current situation perfectly.

The aim of this work is the prediction of current Traffic situation by watching the behavior of available vehicle via a few Mobile GPS. Even if there is at least one available, the system which is based on the way of following figure can detect the condition of current crowding on the targeted road or place.

## 2. System Model

The Figure-1 is the overall architecture of the system. The system can be viewed into three fragments: Client side, Cloud Computation and Cloud Storage. The classification of Transportation mode is completed on Mobile devices. In this classification, the detail identification of Transportation mode, such as walking or bike or driving or bus, is not necessary as a Traffic clue. For this reason, the two types of mode (walking or vehicle) are identified in this work. The

incoming filtered results from individual Mobile device are computed on Cloud by combining the history data of the critical places. The final outcomes are stored on Cloud storage as a history data for future prediction.

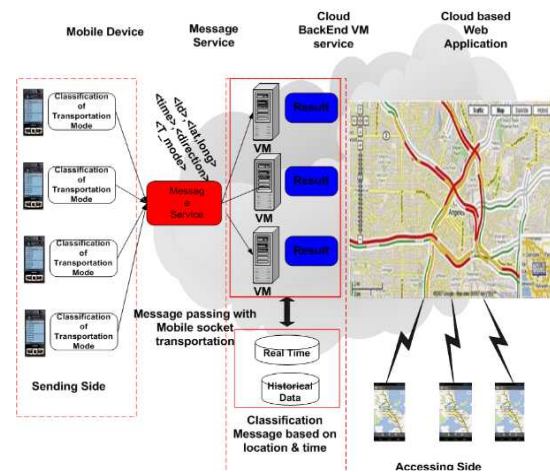


Figure 1. Over All System Architecture

The Traffic Detection System is intended to implement on Google Cloud and the application could be accessed from the Mobile device via Wi-Fi network. In this implementation, Google Application Engine that is the cloud computing service is emphasized on using for building this application. App Engine supports two technology stacks for building web applications: Java and Python. The Java technology stack lets us develop web applications using the Java programming language (or most other languages that compile to Java byte code or have a JVM-based interpreter) and Java web technologies such as servlets and JSPs.

## 3. Related Work

D. Patterson, et- al [1] and L.Liao, et.al [3] use an unsupervised learning technique to detect the transportation mode of a traveler. The transportation modes that are detected in this system include buses, cars and walk. The work in the system is able to predict the traveler's goals, such as trip destination and trip purpose. In addition to GPS data, the system uses historical information about the user. Historical information includes, past user trips and information about where the users parked their cars. In our approach, we only consider the user's transportation mode, motorize or non-motorize, not considers user's history data.

E.Wilson, et- al [2] focus to offer road Traffic state to Mobile user using Mobile network through radio based station and also by analysing a data set provided by a pilot project on the autobahn network in

Southern Germany. V. Jain, et-al [4] also estimates road traffic jams by getting traffic data from CCTV camera installed on roadside and by applying automated image processing mechanism. The difference with my work is using GPS\_enable Mobile phone. V.M. Poulos [5] performed to identify the security and privacy requirements for an Intelligent Transportation System, such as Mobile GPS based Traffic Management System, and propose a solution which meets these requirements by utilizing the existing authentication infrastructure of cellular networks. My work emphasizes to be able to detect the Traffic state even if there is only one vehicle be able to send GPS logs, but not consider the security and privacy of Mobile user.

## 4. Client Process

### 4.1. Preprocessing

**Definition 1.** GPS sensor log. A sensor GPS log  $p_i$  represents data submitted from the GPS sensor embedded on a traveler's mobile device. The format of the report is  $\langle \text{lat}, \text{lon}, \text{t}, \text{v}, \text{h}, \text{acc} \rangle$  where: lat represents the latitude; lon represents longitude; t represents the timestamp of the sensor report; v represents the current ground speed of the device; h represents the direction of travel; and acc represents the accuracy level of the latitude and longitude coordinates. The measurement units of the GPS sensor report attributes are: latitude (lat) and longitude (lon) are in decimal degree; current ground speed (v) is measured in meters per second; direction of travel (h) is specified in degrees counting clockwise from true north; accuracy level (acc) is defined in meters; and time t is in seconds.

**Definition 2.** A GPS trace T is a sequence of GPS sensor reports,  $T = p_0 \rightarrow p_1 \rightarrow \dots \rightarrow p_k$ , where the timestamps in the sequence strictly increase.

**Table 1. GPS sensor log descriptions.**

Notation	Description
P	GPS point which includes latitude, longitude, speed and time
$p_i$	set of GPS point (lat, lon, v, t), $p_i \square P(i=1, 2, \dots, n)$
lat	Latitude
lon	Longitude
t	the timestamp of the sensor report in seconds
v	the current ground speed of the device
h	the direction of travel
acc	the accuracy level of the latitude and longitude coordinates
tm	transportation mode (non walking)
nm	non transportation mode (walking)
R	$\langle \text{id}, \text{lat}, \text{lon}, \text{v}, \text{t}, \text{s}, \text{tm} \rangle$
D	threshold for calculating distance is 0.6
Vt	threshold for calculating average speed is 2

km/s.
-------

### 4.2. Client Process Algorithm

Assume it computes the average speed every 30 seconds.

**Table 2. Client Process Algorithm.**

Client Process Algorithm	
Input :	GPS data= $\langle \text{id}, \text{lat}, \text{lon}, \text{v}, \text{t} \rangle$
Output :	R= $\langle \text{id}, \text{lat}, \text{lon}, \text{v}, \text{t}, \text{s}, \text{tm} \rangle$
Begin	
Step 1.	Receipt $p = \{p_1, p_2, \dots, p_n\}$ within a time window
Step 2.	Average Speed (P.V) = $(\sum_{i=1}^n p_i v) / n$
Step 3.	Filtering speed using Moving Average Filtering Method If (P.V > Vt)
Step 4.	Submit R to backend Server
Else	
Step 5.	Calculate distance (D) within five minutes If (D > threshold)
Step 6.	Submit R to backend Server
Else	
Step 7.	Go to step 1
End	

This algorithm is for computing the filtering process on the mobile device (Client side). Responsively it has to receipt the raw GPS data from the device and then filters out mode of transportation (driving or bus) by calculating these raw GPS data. In other word, it classifies that the mobile device carrier is taking bus or walking. However, there are some problems in analyzing the Transportation mode (walking or not).

However, there are some problems in analyzing the Transportation mode (walking or not). In real world, the running buses have to stop at the any bus stop and at the traffic light. In order to solve this issue, Moving Average Filtering method is applied to smooth the raw and noisy GPS data. This method adjusts the targeted points by replacing the average speed of its neighbors within a span. In equation form, this is written:

$$y[i] = \frac{1}{M} \sum_{j=0}^{M-1} x[i+j] \quad (1)$$

In this equation,  $x[i+j]$  is the number of neighboring data points on either side of  $y[i]$ ,  $y[i]$  is the smoothed output value for  $i$ th data point, and  $M$  is the number of points used in the moving average. Finally, the amount of noise still included in data is computed by taking Mean Square Error equation. In equation form,

$$MSE(X)=E((X-\mu)^2)=\frac{1}{n}\sum_{i=1}^n(x_i-\mu)^2 \quad (2)$$

Besides, the system also measures the distance reached as it take into account this measurement in analyzing Transportation mode. If the final result is Transportation mode (driving or bus), it will send location, speed of device and time to the backend server. Otherwise, the system will begin its process.

### 4.3. Evaluation of Client System

The following Figure 3 and Figure 4 compare the velocity before and after applying Moving Average Filtering method. The method will adjust the fluctuated speed ,Walk ->Vehicle->Walk ->Vehicle->Walk , into a smooth trip, Vehicle->Vehicle->Vehicle->Vehicle->Vehicle. Although a user is travelling by bus, the bus has to stop a moment every bus stop. In this case, sometime the system will recognize itself to be non-Vehicle while its velocity goes under the threshold. This situation is defined as noisy GPS data in this work. Thus, the Figure: 2 is showing that it includes a lot of noise.

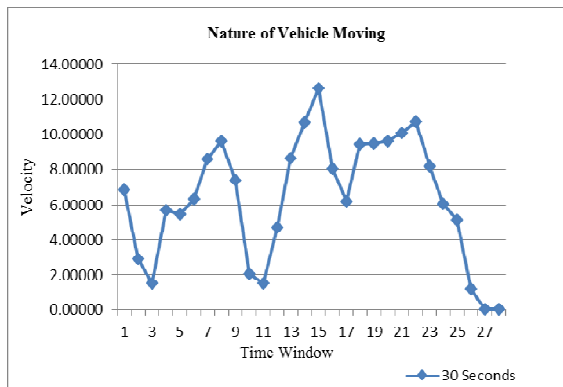


Figure 2. Raw GPS data including Noise

The Figure: 3 shows that it is a clear noise one. These noises are removed out by applying the Moving Average Filtering method by computing the average value of the neighbor and then the noise data will be replaced with this getting average result. By applying this method, the velocity value will be over the threshold. The system will not wrongly determine itself to be non-Vehicle whenever the ground true is a vehicle.

The Table 3 illustrates the estimation errors of analyzing Transportation mode over the data collected for nearly one hours along Pyay road in Yangon, Myanmar.

Table 3. Estimation errors in Analyzing Transportation Mode.

Total Number	Statistics			
	Mean	Median	Standard deviation	MSE
47	5.278858	22.88817	3.941534	0.330547

The accuracy of this part is measured by using MSE equation (2).

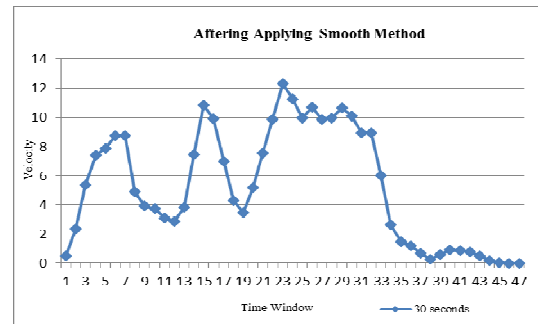
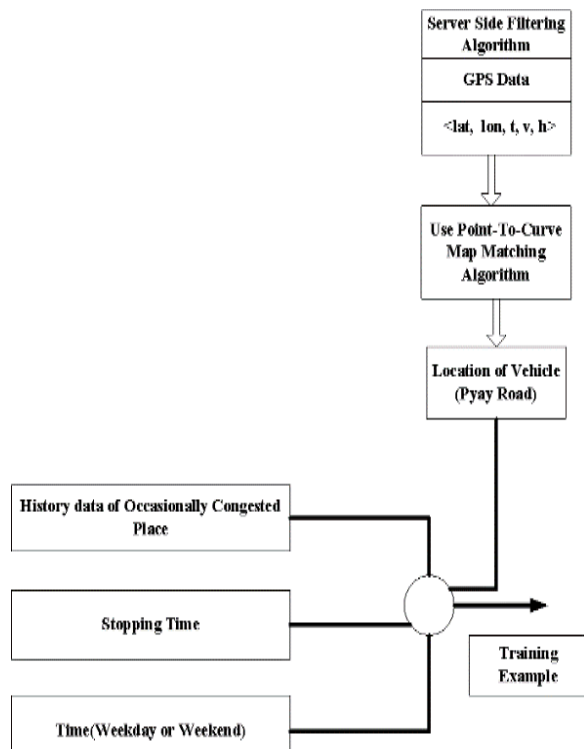


Figure 3. Vehicle's velocity without Noise

## 5. Server Process

### 5.1. Block Diagram of Server Side System

The Figure.4 presents the core modules of Traffic Detection system. Firstly, the system determines the location of GPS filtered log using Point-To-Curve Map Matching algorithm to know in which Traffic zone the vehicle is getting. In Yangon region of Myanmar, the heavy Traffic can be occasionally occurred at Haldan, Myaynigone, Chawdwingone, AD and 8 Mile junction and in a rush hour time, the vehicles are more congested in the evening than in the morning. After detecting the current location, it is continued to learn when, how often and how much the Traffic be occurred upon the history data of that place as a next step.



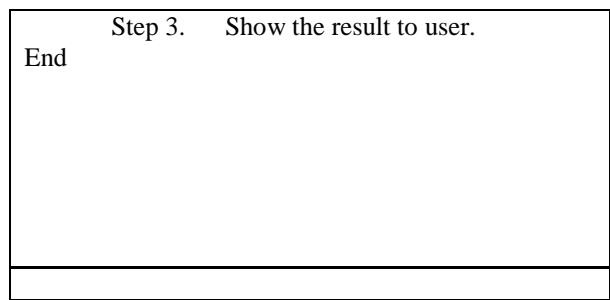
**Figure 4. Traffic Detection Block Diagram**

Afterwards the behavior of vehicle is analyzed through its GPS logs. It means that it calculates whether the velocity is often going under the threshold in the time segmentation and the distance between the segmentations is lower than the distance threshold. Finally, the system combines the information of vehicle and the history data of the critical place and then detects the current Traffic condition by applying Bayes Classifier. The final result will be stored in Cloud storage as a training example for future Traffic detection.

The Table 4 illustrates the process of Traffic Detection algorithm step by step and the final result will be shown on Google Map of android device user.

**Table 4. Traffic Detection Algorithm.**

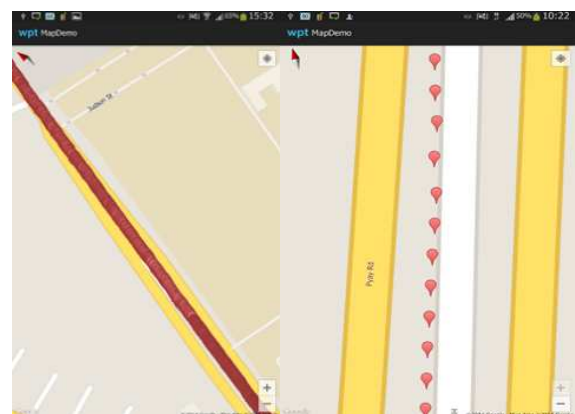
Server Process Algorithm	
Input :	<id,lat,lon,v,t,s,tm>
Output :	traffic status
Begin	
Step 1.	Receipt input <lat,lon,v,t,h> From Server Side Filtering Algorithm
Step 2.	Compute Location of Vehicle using Point-to-Curve Map Matching Algorithm
Step 3.	1. Retrieve the historical data of occasionally congested places 2. Calculate stopping Time 3. Calculate Weekday or Weekend 4. Predict Traffic Status using Naïve Bayes with these data
Step 4.	Store the result as a history data



## 5.2. Data Collection

In the data collection, the required Traffic data for the system has been tracked by implementing an android application to be used as traffic probe in detecting traffic condition. In this collecting data, five smartphones, such as Samsung Galaxy Note8, Sony, Huawei, Galaxy age and Samsung Tab 3 are carried on vehicle from Baham Campus to Hlaing Campus and Insein road at least 3 hours a week including the weekend day over two months.

The following two screen shots present the correct allocation of location data from GPS\_enable Mobile device on the digital map of road network. The needed location data of the targeted road network are extracted from the Google Earth by using Garman tool. Then, these extracted data are segmented and each segment is defined to be 0.01 km long. These segmented GPS points are stored in Cloud storage for matching the geometric relationships between incoming GPS data and a digital map. In this work, the distance of point-to-curve map matching algorithm has been used. Map matching is the process of aligning a sequence of observed user positions with the road network on a digital map.



**Figure 5. Correction Allocation GPS Data Using Distance of Point-To-Curve Map Matching**

The corrected allocation of GPS points has been successfully tested along the Pyay Road in Yangon, Myanmar. The figure.5 shows just the mapping of GPS data to the Geographical data and gives the proof of

the correct mapping of the Geographical data (Latitude and Longitude) extracted from Google Earth segmented into 0.01 km distance with GPS data by using distance of point-to-curve map matching on Google Map. Each segmented is marked by Red Marker.

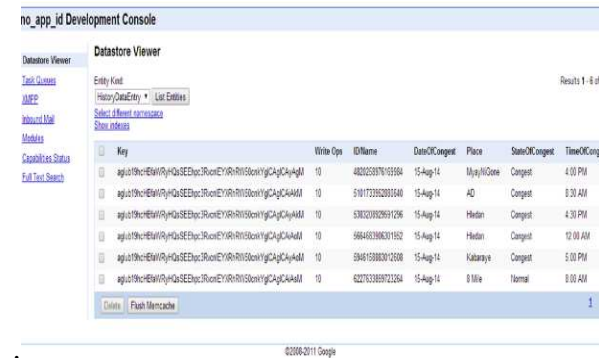
The figure.6 demonstrates the data store viewer of history data of critical congested places in Cloud console. Besides, the information of these places, such as the predefined region, and their attribute, for example, AD junction is one of congested place because there is a school closed to the junction. Hledan junction is sometimes block as the crowd people and vehicle coming to Seingayhar shopping mall and bazaar in this place.

The following figure presents the data format of occasionally congested place stored on the public Cloud storage. It includes the place name, for example AD, on which day congestion is occurred, the state of traffic which means that the traffic is normal or congested and which time congestion is happened, such 12:00 AM or 4:00 PM.

**Table5. Data Format**

PlaceName	DateOfCongested	StateOfCongested	TimeOfCongested
-----------	-----------------	------------------	-----------------

The Figure. 6 gives the data store viewer in Cloud.



**Figure 6. Data store Viewer in Cloud Console**

### 5.3.Evaluation

To predict the condition of Traffic on the targeted road or places, this work applied Naïve Bayes Classifier to predict class membership probabilities such as the probability that a given tuple belongs to a particular class and tested over the training data collected for one month to classify the following instance:

< Speed= Slow, StoppingTime = Long, RegionOfCongestedPlace=Yes, CurrentTime = Normal, TyeOfDay=Weekday, OnRoad=Yes>

This task is to estimate the target value (Yes or No) of the target concept Traffic for this new instance. The target value  $V_{NB}$  is given by

$$V_{NB} = \operatorname{argmax}_{v_j} P(v_j) \quad (3)$$

Where  $V_{NB}$  is the target value, Traffic or No Traffic, and  $\operatorname{argmax}_{v_j}$  that gives maximum values of yes or no.

The input instance is added into the formula as following:

$$V_{NB} = \operatorname{argmax}_{v_j} P(v_j) P(\text{Speed} = \text{Slow} | v_j) P(\text{RegionOfCongestedPlace} = \text{Yes} | v_j) P(\text{CurrentTime} = \text{Normal} | v_j) P(\text{TypeOfDay} = \text{Weekday} | v_j) P(\text{OnRoad} = \text{Yes} | v_j)$$

The probability of the different target values was estimated based on their frequencies over the training examples

$$P(\text{CurrentTime} = \text{Normal} | \text{Yes}) = 0.7$$

$$P(\text{CurrentTime} = \text{Normal} | \text{No}) = 0.3$$

Similarly, the conditional probabilities were be estimated as followed . Those are for CurrentTime=Normal, TypeOfDay=Weekday and OnRoad=Yes.

$$P(\text{CurrentTime} = \text{Normal} | \text{Yes}) = 0.43$$

$$P(\text{CurrentTime} = \text{Normal} | \text{No}) = 0.6$$

$$P(\text{TypeOfDay} = \text{Weekday} | \text{Yes}) = 0.6$$

$$P(\text{TypeOfDay} = \text{Weekday} | \text{No}) = 0.3$$

$$P(\text{OnRoad} = \text{Yes} | \text{Yes}) = 1$$

$$P(\text{OnRoad} = \text{Yes} | \text{No}) = 1$$

Using these probability estimates and similar estimates for the remaining attributes, we calculate  $V_{NB}$  according to Equation (3) as follows.

$$P(\text{Yes})P(\text{Slow}|\text{Yes})P(\text{Short}|\text{Yes})P(\text{Yes}|\text{Yes})P(\text{Normal}|\text{Yes})P(\text{Weekday}|\text{Yes})P(\text{Yes}|\text{Yes}) = 0.1806$$

$$P(\text{No})P(\text{Slow}|\text{No})P(\text{Short}|\text{No})P(\text{Yes}|\text{No})P(\text{Normal}|\text{No})P(\text{Weekday}|\text{No})P(\text{Yes}|\text{No}) = 0.054$$

Thus the naïve Bayes classifier assigns the target value Traffic=Yes to the new instance, based on the probability estimates learned from the training data.



The Traffic result of Bayes classifier in percentage is 0.76 %.

This Traffic percentage generated by Bayes classifier is illustrated in graphical presentation in Figure.7.

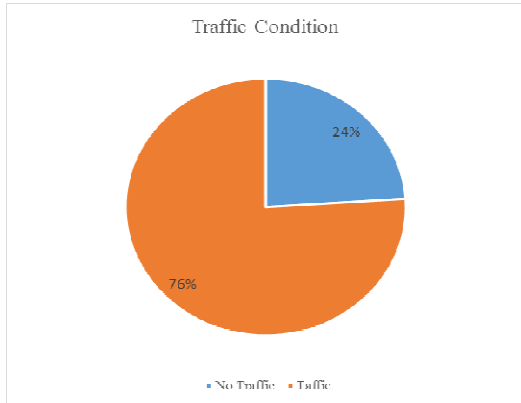


Figure7.Traffic State Generated by Bayes Classifier

The probability of Traffic state not applying Bayes Classifier for the individual place or road was computed by dividing the total number of true outcome with the total number of whole outcome.

$$\text{Probability of Traffic} = \frac{\text{Total number of true outcome}}{\text{Total number of whole outcome}} \quad (4)$$

The Figure.8 compares the Traffic value of AD junction applying Bayes classifier with not applying one over five days. D1, ..., D5 refers to Day. Applying Bayes Classifiers utilizes the history data of the concerning place. Because of that, the result is more improved than not utilizing one.

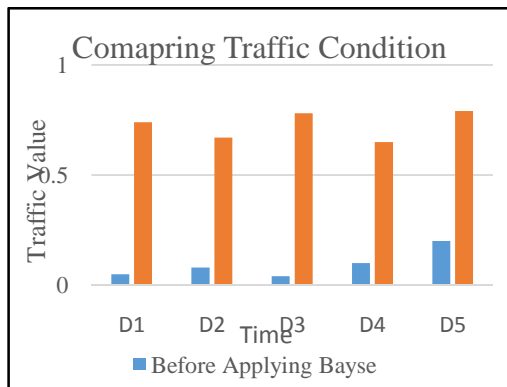


Figure 8. Comparing Traffic State Before and After Applying Bayes Classifier

## 6. Conclusion

Traffic congestion is everyday facing trouble in almost developing and developed countries. This happening makes researchers more interested in and alerts them that Traffic jams is becoming a major challenge. This work approaches to detect the Traffic condition based on analyzing behavior of vehicle by combining the history data as it intends to be able to compute Traffic situation even if the system can get GPS log from only one vehicle at least. Besides, before calculating the Traffic congestion, it filters all invaluable data on the Mobile devices. This means that non-Vehicle mode is not taken into account in detecting Traffic condition. The accuracy of analyzed Transportation mode was measured by utilizing RSE. On Server side, the condition of Traffic is carried out by using Bayes Theorem. The result shows that the situation of Traffic at AD junction is 0.76 %. Future work should investigate in handling and scheduling all incoming Mobile device data for improving processing time on Server side.

## References

- [1] D. Patterson, L. Liao, D. Fox, and H. Kautz, "Inferring High-Level Behavior from Low-Level Sensors", ACM UBICOMP 2003.
- [2] E.Wilson, T.Gould, R.Whittaker and N.Radulovic,"Real-time traffic monitoring using mobilephone data", 2004.
- [3] L. Liao, D. Fox, H. Kautz. Learning and Inferring Transportation Routines. AAAI 2004.
- [4] V.Jain, A.Sharma and L.Subramanian,"Road Traffic Congestion in the Developing World", March 11-12.
- [5] V.M.Poulos, "Security And Privacy in Smartphone Based Intelligent Transportation System",KTH – Royal Institute of TechnologyStockholm, Sweden, 2012.