Travel Time Estimation and Prediction using Mobile Phones: A Cost Effective Method for Developing Countries

Satyakumar, M.¹, Anil, R.¹, and Sivakumar, B.¹

Abstract: Conventional data collection methods lack real time information and involve excessive cost of installation and maintenance. A real-time, low cost travel time data collection system can be developed using mobile phones. This project examines the use of mobile phones for travel time prediction of public transit vehicles and develops a dynamic travel time prediction model. Personnel were employed in public transit vehicles with mobile phones and these mobile phones were tracked continuously. Space information of the mobile phones represents the position of the buses and movement pattern of these mobile phones in turn represents the movement pattern of the public buses. The starting and arrival time at sections obtained from the cellular database were used to get the travel time and speed. Results obtained were statistically significant and it shows that use of mobile phone for travel time data collection is a low cost data collection technique for Indian cities.

Keywords: Resource Tracking and Management Service (RTMS), mobile phone tracking, real time traffic data collection.

Introduction

With the quickly increasing number of vehicles in the recent years, the urban traffic is facing serious problems. Because of the limitation of space, it is not always possible to construct new roads. But, with the innovation of new technologies in computer, communication, sensors etc., many applications can be developed to improve traffic situations by utilizing the existing infrastructure itself. One of the main aspects of traffic management is the incident management. Speed and travel time measurements on real-time are the most extensively used traffic data to monitor incidents. On the other hand, road users rely on travel time information for journey planning and route choice decisions. More recently Advanced Traveler Information Systems (ATIS) are being developed to provide real time information to roadway users to help them in making various decisions, such as: route choice, and departure time choice. Sutandi [1] had evaluated that Variable Message Sign (VMS) has significant impact on traffic performance measures in large cities with higher number of alternative roads in developing countries. Application of VMS as one of Intelligent Transportation Systems (ITS) technologies in congested roads in developing country can potentially ease congestion problem in the cities. All these systems rely fundamentally on the availability of real-time data on current traffic conditions.

The current state-of-practice for data collection regarding traffic speeds relies mainly on local detectors that measure the speed at a specific point along the roadway. One of the most widely used technologies for this purpose is magnetic loop detectors, installed under the roadway surface. Due to the cost of installation and maintenance of local detectors, they are typically installed only on a relatively small portion of the roadway system, resulting in a limited coverage of the entire transportation network. In addition, in an urban environment there are many traffic interruptions, particularly at intersections, which cause delays that are not easily depicted by simply measuring speeds at any specific point along the road. Another type of detector is video observation, by which the general traffic situations within a given area are directly shown, but here also factors like velocity of vehicles, flow rate etc., still cannot be estimated quickly. An alternative approach is to measure the travel times of vehicles along a certain route or route segment. When a small number of occasional measurements are needed, dedicated floating vehicles can be used. Equipping floating vehicles with Global Positioning System (GPS) can improve the accuracy of the measurements. But data from a fleet of equipped vehicles, typically limited in its size, is also limited in its coverage, especially when real time data is needed. Furthermore, specific fleets often have specific travel patterns that are not necessarily representative of the entire population.

¹ Traffic and Transportation Engineering Division, Department of Civil Engineering, College of Engineering Trivandrum, Kerala, INDIA. Email: bsivakumar007@yahoo.co.in

Note: Discussion is expected before June, 1st 2014, and will be published in the “Civil Engineering Dimension” volume 16, number 2, September 2014.

Received 06 May 2013; revised 06 November 2013; accepted 12 December 2013
Recent advances in wireless communication technologies provide the capability to determine the location of wireless communication devices such as mobile phones. Cellular service system contains information about the locations of its users over time. This capability provides an opportunity to acquire large quantities of traffic data without the need to deploy a large and costly network of dedicated traffic sensors. Vehicle probes using mobile phones can be considered a promising technology for generating real-time travel information. In recent years, mobile communications have been more and more popular and the technologies of mobile communication and the mobile phone equipment itself have also been improved a lot. The ever growing market of mobile phone users have the great potential to be used as probes for dynamic tracking of vehicles in the transportation network.

In this study, an attempt has been made to explore this potential of mobile phones, for travel time data collection and to predict the arrival time of public transit units. In order to reduce congestion and other traffic related problems, the most efficient traffic management measure is to make the Public Transportation System (PTS) attractive and reliable so that the road users should switch from personal mode to the PTS. For making the PTS attractive, there are many methods like bus priority measures and Intelligent Transportion System (ITS) applications for better operations and management. One of the most widely used ITS applications is providing the bus arrival time in advance. The problem of predicting travel time on urban roads has been a topic of research for a long time in many of the developed countries, but it was rarely been a research topic in Indian context. Only very few studies have been conducted in India so far. In the following section, current literatures on this topic are reviewed.

Mobile Communication System

Mobile phones in a network are identified based on Cell-ID. As the position of every Base Transceiver Station (BTS) is fixed, the location of a mobile phone can be estimated through the Cell-ID information of the BTS which is serving it [2]. Physically, BTS is the access point of the mobile handset to the network (Figure 1) [3]. The area covered by one BTS is called a Cell. The Base Station Controllers (BSC) is a device that controls multiple BTS. It handles allocation of radio channels, frequency administration, power and signal measurements from the Mobile Station. The heart of the Global System for Mobile communications (GSM) network is the Mobile Switching Centre (MSC), which handles multiple BSC and also interfaces with other MSC.

![GSM Network – Schematic Diagram](image)

Location Management from GSM network is possible through a system of databases, which contain data about subscribers, including current location of the mobile phones. During a signal exchange between mobile and the tower, the location management processes determine the position of the mobile at the Cell ID level. The operator knows the coordinates of each cell site and can provide an accurate position of the connected mobile handset through Received Signal Strength (RSS) method or Triangulation method.

This location technique can be used to obtain location and travel time information. For instance, if the person with mobile phone in a bus can be considered as a ‘probe vehicle’, the mobile locations then become the locations of the bus, by which information such as speed, travel time, distance travelled, etc., can be estimated. Every time, latitude, longitude and time are obtained and transmitted to the location database established in a server in the information processing center. Then, through Geographical Information System (GIS) integration the actual traffic situations on the road network within an area can be directly shown in maps which will give a real time display, as the database is always updated by the new data from mobile phones.

Minor location precision errors of about 200-250 m existing in the system at present, becomes less important when traffic is constrained for a whole highway network connecting different cities. Other instruments like GPS, providing greater precision do however exist; but they either require modification of the network (installation of ancillary devices to measure the signal, reception angle, signal level and so on) or involve alteration of the terminals (GPS, clocks to record reception times and so on) to collect the data. In view of the costs involved in such supplementary modifications and the fact that the traffic data inferred need to be representative; that is, the sample employed should be as large as possible, positioning provided by a GSM network may be more appropriate for traffic surveys. Some of the advantages of GSM technology include large data sample, large aerial coverage, no installation of additional devices either in vehicles or in the network, real time data generation etc.
Location Based Services (LBS)

Mobile phone service providers offer Location Based Service (LBS) to its subscribers which give the subscriber information on their vehicles, resources, fleet management, etc. These value-added applications available to the users are based on the geographical location of the mobile phones. LBS can be used to provide real-time positioning of a mobile device in a vehicle in the transportation network. It allows simultaneous tracking of location and status of vehicles using a standard Subscriber Identification Module (SIM) in a mobile phone or other mobile devices. The phone acts as an intermediate tool to access data on its location and will position it on a map online. As well as tracking vehicles, the travel time measured with mobile instruments can be used for evaluating the level of service of transport systems in urban area.

Literature Review

Several attempts can be found in the past few years to adapt the mobile phone technology for transportation planning and operations purposes. Projects regarding feasibility and field-tests of mobile phone location-based ITSs have been developed in Rome [4], Israel [5], Spain [6], Bangkok [7] and China [8]. Most of the projects and studies have proven the feasibility of this real-time traffic data collection using mobile phone technology.

Vehicular traffic status and instantaneous location of buses and taxis can be described through the anonymous monitoring of the mobile phone movement [4]. In a study conducted in Israel [5] it was found that the speed and travel time measurements from mobile phone based system possess a good match with those obtained by dual magnetic loop detectors. The system used in this study focuses on handover events at which control of a presently used phone is handed over from one cell to another. For every handover, the system associates a geographical footprint, in which the phone has been most likely when the handover occurred. Results show that GSM mobile phone network can also be used to derive Origin-Destination (OD) data [6]. The flow of mobile phones in a cell-phone network is measured and correlated to traffic flow. This method has great potential and the results inferred are much more cost-effective than those generated with traditional techniques. Cell Dwell Time (CDT) has been used by Pattara-atikom and Peachavanish [7] to estimate the degree of traffic congestion. CDT is the duration that a mobile phone is registered to a base station before handing off to another base station. The results showed promising performance of congestion estimation with accuracy of 79.43%, precision ranging from 73.53% to 85.19% and mean square error of 0.44. The short message function and locating function of the mobile communication system were also been used to get various parameters like travel time, OD Matrix, trip purposes, and mode of travel [8]. It was also possible to estimate congestion level along a study stretch using mobile phone travel time data in a study of estimating man-hours lost due to congestion [9].

Present Study

Bharat Sanchar Nigam Limited (BSNL), the leading mobile phone service provider in India provides LBS such as Resource Tracking and Management Service (RTMS), People Finder Service (PFS), Navigation, etc, to its customers. RTMS is designed for corporate customers to manage and track vehicles, mobile employees and assets over the internet through a web interface. It uses online mapping facilities to track the vehicles and it can also track different vehicles simultaneously. Details such as location and path, time of starting, time of arrival, and distance travelled of the tracked vehicle, based on the user defined parameters such as landmarks and the tracking time interval, will be available over a secured internet web interface to the subscriber. From these details, travel time and speed of the vehicles can be calculated. As a pilot study, RTMS facility of BSNL is subscribed for 14 SIM and for the survey to examine the emerging opportunities of collecting traffic information using mobile phones as traffic probes. The following information on congestion can also be extracted from the real time travel time measurements obtained through RTMS:

- Spatial location of the congestion in the city area: which roads are experiencing the congestion and where exactly are the starting and ending points of the congestion.
- Temporal characteristics of the congestion: in what times of the day the congestion is experienced most and in what days of the week.
- Unusual congestion: occurrence of accidents or any other incidents may be easily identified by an unusual congestion experienced in the route through an unusual increase in travel time in that particular stretch.

In the present study, times of travel were collected for a typical week using mobile phones with RTMS activated SIMs. This study aims at preparing a travel time prediction model for public transit vehicles. The 10.2 km, East Fort – Sreekaryam stretch of National Highway 47, which is a major corridor in Trivandrum city was selected for the study.
Method

Kerala State Road Transport Corporation (KSRTC) is the largest public sector undertaking in the State of Kerala with a fleet strength of 5,684 public transit buses and running 1,560,000 km carrying 3,400,000 to 3,700,000 people a day in Kerala and in the neighbouring States. Personnel were sent simultaneously at 10-15 minute interval along the study route in KSRTC buses in the morning peak hours (8.00 am to 11.00 am) from both ends of the route with the mobile handset containing the RTMS tracking facility activated SIM. During their travel they were continuously tracked at an interval of five minutes and the location information were stored in the BSNL database. At the same time the personnel recorded their trip details such as time of start and time of end of the trip, for checking the accuracy of LBS data. Exact distance of the study corridor was recorded by travelling in a car.

The information obtained from the RTMS report generated in the user web-interface of BSNL-RTMS are address, longitude, and latitude of the predefined landmark along with the arrival time of the mobile subscriber near that landmark. The RTMS data were extracted to get the longitude and latitude of the locations where the mobile phones were tracked and the time of arrival of the mobile phones at these locations. Entering these longitude and latitude values in Google Earth [10] online, these locations where identified. Time of travel between the end points was calculated from the time of arrival details at these locations and average journey speed of public buses was obtained. Tracking has been done for four consecutive days of a week and Kalman Filter Algorithm [11-14] was developed to predict the travel time for the fifth day, taking four day’s tracking details as history. A detailed analysis is given in the following section.

Results and Analysis

Travel times of a given road stretch may be used to identify the service quality provided as well as their variations. For example, Figure 2 and 3 show the travel time variation along the Sreekaryam – East Fort and East Fort – Sreekaryam routes respectively, for different departure time from 0800 to 1900 on a day.

From Figure 2, it can be observed that for the morning peak hours travel time is high for the period 0900 to 1000, indicating a severe congestion along the route during that period. For the trips starting at 0940 and 0950 the travel time reaches a peak value of 65 minutes. The travel time eventually reduces to lower values from 1020 to 1330 and reaches 30 minutes at 1330. So this travel time of 30 minutes is taken as the minimum travel time along the stretch. This minimum travel time is taken as the benchmark for calculating delay and average desirable journey speed of buses and getting the congestion information in the further analysis.

A secondary pattern of variation in travel time occurs during evening peak after 1700. The worst condition occurs to the trips started at 1700 and 1735. Maximum travel time of 80 minutes is taken during these trips.

These increased travel times were perhaps a result of enhanced traffic along some of the bottle necks of the route. These travel time values gives a real picture of the traffic conditions along the study route. In addition to these high changes in travel time, small variations occur during noon off-peak hours due to some local incidents.

At the same time in the opposite direction, we can see that the congestion is high during the period 0800 to 0915 and 1715 to 1830 (Figure 3). Peak travel times were experienced for the trips started at 0830 and 1735. It can be seen in Figure 3 there is an abrupt change in travel time for the trip started at 0830, this may be because of a short-period traffic block due to some incidence like breakdown of a vehicle or so.
Thus, information on traffic condition was available in real-time through the tracking of mobile phones. This information of extended travel time will be obtained when the mobile is tracked for the next instant. At that instant, the actual distance travelled by the mobile phone will be less than the proposed distance. Thus mobile tracking makes travel time information available in real-time and hence provide real-time information of traffic situations along the route.

Speed is another parameter, which has been used widely to assess the quality of service of a road depending on the demand it supports, especially in urban environments. Speed is normally measured at a point or short section of a road to determine the rate of motion of vehicles passing through it. In this study an attempt is made to get an idea of variations in the average journey speed of public transit vehicles using the real-time data obtained through RTMS. From the RTMS data, speed is calculated using the average travel time and length for the roadway segment.

Figure 4 shows average journey speed variation of public buses with respect to the time of departure for the South bound flow. Maximum desirable journey speed of buses was obtained as 20-40 km/h, by considering off-peak travel time as the minimum journey time. Here again we can see that the congestion is at its peak for the trips starting during the interval 0900-1000 and 1700-1800 and high congestion occurs for the trips started at 0940, 0950, 1716 and 1735.

The actual average journey speed for buses in the South bond direction, without considering the stopping time, is only less than 10 km/h during the morning peak, while in the North bound direction the situation was worse for the trips starting at 1715 and 1735 (Figure 5). But the trips started at 1330 and 1455 in the South bound direction and trips started at 1020, 1030 and 1100 in the North bound directions possess the desirable average journey speed of 20-40 km/h.

Figure 5. Average Journey Speed Variation with the Time of Departure from East Fort to Sreekaryam Direction (Northbound Trip)

Travel Time Prediction Model

In this study Kalman Filter Algorithm [11 - 14] is used to predict the arrival time of buses along the study route. Models were developed using four days travel time as history. Using the model, travel time for Friday morning peak was predicted with the help of morning peaks of Monday to Thursday.

Kalman Filter Algorithm is a filter operation usually applied in the field of Electronics Engineering to recover the disturbed signals by analyzing the historical pattern. The accuracy of the filter operation increases with the increase in the number of iterations. As the travel time at any instant is highly correlated with the travel time of preceding identical vehicle, the last travel time of a bus provides a fair indication of how long it will take for the next bus to drive through the same part of the route. Using this concept a Kalman Filter Algorithm can be developed to predict the arrival time of public buses from the historical arrival data.

Generally, the Kalman Filter Algorithm for bus link running time has the following structure:

\[ g_{t+1} = (e_t + \text{VAR}_{\text{avg}})/(e_t + 2\times\text{VAR}_{\text{avg}}) \]
\[ a_{t+1} = 1 - g_{t+1} \]
\[ e_{t+1} = \text{VAR}_{\text{avg}} \times g_{t+1} \]
\[ P_{t+1} = a_{t+1} \times P_t + g_{t+1} \times \text{art}_{t+(k-1)} \]

where: \( g \) = the filter gain, \( a \) = loop gain, \( e \) = filter error, \( P \) = prediction,

\( \text{art}_t \) = actual running time of the previous bus at instant ‘t’,
\( \text{art}_{t+(k-1)} \) = actual running time of the bus at instant \( t+1 \)’ on the previous day,
\( \text{art}_{t+(k-1)} \) = actual running time of the bus at instant \( t+1 \)’ two days ago,
\( \text{art}_{t+(k-1)} \) = actual running time of the bus at instant \( t+1 \)’ three days ago,
\( \text{art}_{t+(k-1)} \) = actual running time of the bus at instant \( t+1 \)’ four days ago,
VARAvg = Average of the last four days variances
= \frac{(\Delta_1+\Delta_2+\Delta_3+\Delta_4)}{4} \quad (5)

\Delta_1 = [\text{art}_{t+k+1} - \text{avg}(\text{art})]^2 \quad (6)
\Delta_2 = [\text{art}_{t+k+1} - \text{avg}(\text{art})]^2 \quad (7)
\Delta_3 = [\text{art}_{t+k+1} - \text{avg}(\text{art})]^2 \quad (8)
\Delta_4 = [\text{art}_{t+k+1} - \text{avg}(\text{art})]^2 \quad (9)

Now it is easy to implement the Kalman Filter Algorithm using these equations applying in the order given (Table 1). Obviously in real-world application this algorithm can make use of longer range of historical data.

Model Performance Evaluation

In order to assess the predictive performance of the models, the predicted values are compared with actual values through a statistical measure. Mean Absolute Percentage Error (MAPE) is the statistical measure used here to test the model performance. MAPE is given by,

\[
\text{MAPE} = \frac{1}{n} \sum \frac{|\text{Field data} - \text{Simulated data}|}{\text{Field data}} \times 100
\]

(5) MAPE value of 0-10% = Very Good, 10-20% = Good, 20-30% = Satisfactory. Detailed calculation for South bound flow is given in Table 1. The model had shown MAPE values in the range 10-20% indicating that this model performs very well.

Summary and Conclusion

This paper describes potentials of a new cost effective method based on passive data collection from a mobile phone provider for real-time traffic data collection. Given that alternative options to estimate real-time congestion characteristics are not available; this method could fulfill the requirements at a reasonable accuracy level and timely effort. Since public buses had been utilized as probes for data collection, no cost was incurred in providing infrastructure facilities. In this method either the driver or the ticket collector in the public transit unit can act as a probe thus reducing the man power requirement, and hence becomes more economical compared to the traditional methods. Moreover, the results obtained were more reliable as there were minimum human interventions and hence reduced the chance of errors.

The method experimented in this study works well for measuring travel time of a given corridor. In a micro level analysis, the expected arrival time at various bus stops can also be predicted using the historical data of previous days. Thus the prediction model used in this study can provide continuous information on the expected arrival time of buses at downstream stops and hence information on the expected deviations from the actual schedule. This facility enables the authorities to assess in real-time, transit stop-based control actions to avoid such

<p>| Table 1. Kalman Filter Algorithm Calculations for Bus Running Time (Southbound Trips) |
|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|</p>
<table>
<thead>
<tr>
<th>Time of Departure (1)</th>
<th>Actual Travel Time Taken</th>
<th>Predicted Time (14)</th>
<th>Observed Time (15)</th>
<th>Absolute Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before 8:00am</td>
<td>30</td>
<td>30.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>8:00-8:15</td>
<td>45</td>
<td>40.0</td>
<td>35.0</td>
<td>35.0</td>
</tr>
<tr>
<td>8:15-8:30</td>
<td>55</td>
<td>40.0</td>
<td>40.0</td>
<td>40.0</td>
</tr>
<tr>
<td>8:30-8:45</td>
<td>55</td>
<td>42.5</td>
<td>40.0</td>
<td>45.0</td>
</tr>
<tr>
<td>8:45-9:00</td>
<td>55</td>
<td>45.0</td>
<td>40.0</td>
<td>55.0</td>
</tr>
<tr>
<td>9:00-9:15</td>
<td>55</td>
<td>50.0</td>
<td>57.5</td>
<td>57.5</td>
</tr>
<tr>
<td>9:15-9:30</td>
<td>50</td>
<td>60.0</td>
<td>55.0</td>
<td>60.0</td>
</tr>
<tr>
<td>9:30-9:45</td>
<td>40</td>
<td>60.0</td>
<td>52.5</td>
<td>65.0</td>
</tr>
<tr>
<td>9:45-10:00</td>
<td>40</td>
<td>50.0</td>
<td>50.0</td>
<td>65.0</td>
</tr>
<tr>
<td>10:00-10:15</td>
<td>40</td>
<td>50.0</td>
<td>45.0</td>
<td>62.5</td>
</tr>
<tr>
<td>10:15-10:30</td>
<td>40</td>
<td>47.5</td>
<td>42.5</td>
<td>60.0</td>
</tr>
<tr>
<td>10:30-10:45</td>
<td>45</td>
<td>45.0</td>
<td>40.0</td>
<td>60.0</td>
</tr>
<tr>
<td>10:45-11:00</td>
<td>50</td>
<td>40.0</td>
<td>35.0</td>
<td>40.0</td>
</tr>
</tbody>
</table>

MAPE = \frac{1}{n} \sum_{i=1}^{n} e_i = 8.7
deviations before their occurrence, hence allowing for proactive control, as opposed to the traditional reactive control which attempts to recover the schedule after deviations occur.

As the present study is based on a system where by the output is originally designed for mere tracking of commercial vehicles of corporate customers, there are some challenges that need to be overcome such as identification of the phones that are in-vehicle or not in-vehicle. In future, it may be possible to generate models for trips grouped by day, time of the day, etc. Furthermore, as the ITS deployment continues, the models could be expanded to include traffic condition variables, such as congestion and incidents, that can be automatically generated by these systems. From this study, it could conclude that there is a wide scope for the traffic data collection using mobile phone networks in coming future.

Acknowledgment

The authors wish to thank Ministry of Urban Development (MoUD), Govt. of India and Centre of Excellence in Urban Transport, Department of Civil Engineering, IIT Madras for financial assistance and guidance provided by them for the successful completion of this work. Authors also express sincere gratitude to staff members, BSNL, Trivandrum for the extreme help rendered by them, without which this work could not have been materialized.

References