NEW SNAPSHOT GENERATION PROTOCOL FOR TRAVEL TIME ESTIMATION IN A CONNECTED VEHICLE ENVIRONMENT

CHEN CHEN, M.S.
Graduate Student,
Department of Civil and Environmental Engineering,
University of Missouri, E2509 Lafferre Hall, Columbia MO 65211

JALIL KIANFAR, PH.D.
Transportation Analyst,
Crawford, Bunte, Brammeier (CBB),
1830 Craig Park Court, Suite 209, St. Louis, Missouri 63146

PRAVEEN EDARA, PH.D., P.E.*
Associate Professor,
Department of Civil and Environmental Engineering,
University of Missouri, C2640 Lafferre Hall, Columbia MO 65211
Email: edarap@missouri.edu, Ph: (573) 882 1900, Fax: (573) 882 4784

Paper number: 14-1724
Revised Submission Date: November 15, 2013
Word Count: 7,013 (4,513 text plus 2500 (5 figures and 5 tables))
* Corresponding author
ABSTRACT

Connected vehicle technology offers great potential in improving the safety and mobility of a transportation system. Probe data collection is one feature of connected vehicle technology, where vehicles collect information such as their location and speed. Probe data could be used to support various traffic management and traveler information applications. In this paper, a novel protocol called the R² protocol is presented for collecting probe data in a connected vehicle environment. The core principle of R² protocol is to only collect vehicle snapshots when a significant change occurs in vehicle speed. Data from a connected vehicle simulation test-bed and a real-world test-bed in Michigan were used to evaluate the proposed protocol. Average speed method and a method based on reconstructing vehicles’ time-speed plots were used to estimate link travel time. Linear regression, cubic spline, and piecewise cubic Hermite interpolation were applied to reconstruct time-speed plots. The proposed R² protocol was compared with three existing protocols, the Fixed 2 second, Fixed 4 second, and SAE J2735 protocols. The results from the simulation test-bed indicated that not only did the R² protocol outperform the three protocols in terms of the error measure but it also required fewer snapshots to achieve that value. The number of snapshots recorded by the R² protocol was 30%, 26% and 4% lower than the Fixed 2 second, Fixed 4 second, and SAE J2735 protocol, respectively. The Michigan test-bed case study showed that the R² protocol had smaller error and needed 11% fewer snapshots than the SAE J2735 protocol.
INTRODUCTION

Connected vehicle technologies offer new ways to advance the safety and efficiency of a transportation system. Previous research has demonstrated the benefits of these technologies in designing lane change advisory systems (1), intersection signal control and queue detection (2-3), transit signal priority (4), travel time estimation (5-6), work zone traffic control (7), sustainability (8), and pavement monitoring (9). Studies have investigated the connected vehicle applications within the environment outlined by the SAE J2735 standard (1-8). The connected vehicle environment in the SAE J2735 standard (10) consists of three major elements: On-Board Unit (OBU), Roadside Equipment (RSE), and Dedicated Short-Range Communication (DSRC). OBUs are installed in vehicles and record vehicle activity data; RSEs are installed at intersections, interchanges and other locations to provide communication interface to the vehicles. When a vehicle enters a RSE coverage area the information stored in an OBU is transmitted to the RSE. DSRC is a wireless communication channel specifically designed to support automotive and transportation applications.

In the SAE J2735 standard, several message sets are defined to support a wide range of intelligent transportation system (ITS) applications. Vehicle probe data is an example of such message sets; data provided by this message set could be used in traffic management and traveler information systems. In this message set OBUs record three types of snapshots: periodic snapshots, start and stop snapshots, and event triggered snapshots. Periodic snapshots are recorded at regular intervals. When vehicle speed is greater than or equal to 60 mph, periodic snapshots are recorded at 20-second intervals. When vehicle speed is less than or equal to 20 mph, snapshots are recorded at 4-second intervals. For speeds between 20 mph and 60 mph, a linear approximation between 4 seconds and 20 seconds is used to identify the snapshot interval. Stop snapshots are recorded when, 1) a vehicle does not move (i.e., speed=0) for five seconds, and 2) when there is no record of vehicle stopping within a 15-second interval. When a stop snapshot is recorded, periodic snapshot is no longer recorded. A start snapshot is recorded when the vehicle speed exceeds 10 mph. Event triggered snapshots are recorded when vehicle status elements change. Airbag activation is an example of an event-triggered snapshot.

A probe data collection protocol with fixed-interval snapshots and preferably short intervals has been also recommended as an alternative to SAE J2735 probe data protocol (5). Collecting probe data in short intervals increases the sampling rate and comes with the cost of larger data size and redundant samples. This paper presents a novel probe data collection protocol called the R^2 protocol that only records snapshots when a significant change occurs in vehicle speed. The snapshots are collected in such a way that a vehicle’s time-speed plot could be reconstructed from the collected snapshots with a predefined level of accuracy (for example, 95%). The developed protocol not only reduces the sample size of probe data but also improves the quality of data by capturing all significant changes in vehicle speed. There are several benefits associated with eliminating the redundant snapshots and thus reducing the sample size. First, a limited number of snapshots can be stored in a vehicle OBU buffer. If OBU buffer capacity is reached before the vehicle comes within communication range of a RSE, older snapshots are replaced by newer snapshots resulting in loss of probe data collected in earlier segment of the travel itinerary; eliminating redundant snapshots would allow OBU buffer to store snapshots for a longer period of time before snapshots are replaced with newer snapshots thereby allowing speed monitoring over a longer segment of vehicle travel itinerary. Second, RSE units can only communicate with a limited number of OBUs simultaneously; reducing the sample size would also reduce the communication load, and would result in shorter data transmission time from OBU to RSE. The smaller sizes would
allow the RSE to serve more vehicles. Third, traffic management centers could also benefit from smaller storage and processing needs of smaller samples of connected vehicle probe data.

Data from a connected vehicle simulation test-bed previously developed by the authors (6) and data from a real-world connected vehicle test-bed in Michigan were used to evaluate the $R^2$ protocol performance. Three other protocols: SAE J2735, Fixed 2-second (11), and Fixed 4-second (11) interval were also evaluated and compared with the $R^2$ protocol performance. Arterial travel times were estimated using the collected probe data. Two travel time estimation methods, the average speed-based travel time method (6), and a new method based on reconstructing vehicle time-speed plots were used for the evaluation. For the new travel time estimation method, time-speed plots were reconstructed using interpolation. Three interpolation methods: linear, cubic spline, and piecewise cubic Hermite, were investigated.

LITERATURE REVIEW

This section presents a review of the literature in two areas. First, the existing literature on snapshot protocol use in a connected vehicle environment is discussed. Second, the state of the art in travel time estimation methods is presented.

Connected Vehicle Probe Data

Shladover and Kuhn (12) investigated the quality of connected vehicle probe data for adaptive signal control, incident detection and weather condition monitoring systems. They used VISSIM to simulate approximately 10 km of SR-82 in Palo Alto and Mountain View in California. Vehicle trajectories were recorded and processed based on the SAE J2735 standard. They concluded the data collected based on current probe data protocol provides an acceptable representation of normal traffic conditions assuming 1 to 2 minute data latency is acceptable. Park et al. (13) developed a connected vehicle simulation test-bed by integrating VISSIM traffic simulation software and NCTUns communication simulator. SAE J2735 message sets were implemented in the test-bed to facilitate evaluation of various connected vehicle applications. Dion et al. (11) developed a virtual test-bed for connected vehicles and evaluated the probe data generated in connected vehicle environment. They simulated the USDOT test-bed in Michigan in Paramics traffic simulation software and followed the SAE J2735 standard to collect data. In another study Dion et al. (5) evaluated the issues related to usability of connected vehicle generated probe data based on current standards. They recommended several improvements to current probe data protocols. In terms of snapshot intervals, they recommended fixed-interval snapshots and preferably short intervals. They also advocated for a protocol that generates snapshots while vehicles are stopped.

Connected Vehicle Travel Time Estimation

Significant efforts have been directed toward improving link travel time estimation methods. Such methods include tracking cell phones on the network (14-15), tracking Bluetooth devices on the roadway (16-18), using toll collection data (19), automatic vehicle identification (AVI) (20), and vehicle signature analysis (21-22). However, very few studies have investigated the application of connected vehicle data for travel time estimation. Oh et al. (23) developed “ubiquitous probe vehicle surveillance system (UBIPROSS)” that uses GPS and V2V communication to collect probe data. Two different methods were used for travel time estimation: 1) length of the link was divided by average speed of all uploaded snapshots of vehicles on the link, 2) space mean speed was used to calculate link travel time.
Rim et al. (24) developed a travel time estimation method that uses vehicles’ speed and coordinates provided by V2V and V2I to estimate lane-level travel times. The structure of model was based on concept of dynamically defined links and nodes. Nodes were defined with respect to location of incidents and the area impacted by the incident. Speed observations were categorized into different groups and weighted average of vehicles’ speed was used to calculate link travel time. This paper takes a different approach to travel time estimation. Time-speed plot of vehicles on a link are reconstructed and link travel time is estimated using initial and final speed data of a vehicle on the link. Although this paper is the first attempt to reconstruct time-speed plot of vehicles from connected vehicle data, Sun et al. (25) suggested a “piecewise truncated quadratic” model to estimate speed profile between point detectors, such as loop detectors. Piecewise quadratic interpolation was used to create the speed profile of the two road segments between three successive detectors. Lower and upper bounds were imposed to the outputs of piecewise quadratic interpolation to better represent congested flow conditions and avoid unrealistic speed estimates. The reconstruction of time-speed plots in this study has some similarities with the vehicle trajectory analysis used for shock wave analysis. Lu and Skabardonis (26) developed a method for estimation of shockwave speed using vehicle trajectory data. Izadpanah et al. (27) used linear regression to identify significant changes in vehicle trajectory data. The “change point” in a vehicle trajectory was determined using an optimization program. Cheng et al. (28) propose a method that identifies “critical points” in a vehicle trajectory as instances where a change in vehicle movement occurred. The proposed method was able to remove redundant trajectory data.

22 METHODOLOGY

In the existing probe data protocols, snapshots are recorded in fixed intervals such as 4 seconds or in a predefined range of intervals such as between 4 seconds and 20 seconds depending on the vehicle speed. This section presents the $R^2$ snapshot data collection protocol in which snapshots are collected only when a significant change occurs in the vehicle speed. The travel time estimation method based on speed plot reconstruction is also presented in this section.

29 $R^2$ Protocol for Probe Data

The $R^2$ protocol is designed to collect only a sample of snapshots sufficient enough to construct an accurate time-speed plot for each vehicle. The $R^2$ protocol is analogous to feature detection in computer vision and image processing where a rapid change point is defined as a corner. A corner can be detected by looking for high levels of curvature in the image gradient. The intent behind $R^2$ protocol is to find “corner” points in time-speed plots. By capturing only the significant changes in vehicle speed the number of redundant snapshots collected will be minimized. Although the feature detection approach has been used in shock wave analysis as described earlier, there are some key differences between the proposed $R^2$ protocol and the methods used in previous research. The $R^2$ protocol differs from the linear regression approach used by Izadpanah et al. (27) since only up to four snapshots are temporarily stored in the OBU buffer while there is no limit for the timespan between temporarily stored snapshots. The $R^2$ protocol also differs from the method used by Cheng et al. (28) since it guarantees that the collected snapshots when used to reconstruct time-speed plot will produce the desired $R^2$ value.

The final output of $R^2$ protocol is a time-speed plot that could be divided into virtual segments. Figure 1 provides an example of a time-speed plot recorded under the $R^2$ protocol.
FIGURE 1  Example of Time-Speed Plot Recorded under R² Protocol

In this figure blue lines with arrows are used to show virtual segments of the time-speed plot. Each segment starts and ends with a snapshot. The coefficient of determination between a linear fit connecting the start and end points of a segment, and the ground truth speed observations is greater than or equal to the predefined R² threshold.

In the R² protocol, the end snapshot of segment \( i \) serves as the beginning snapshot of the segment \( (i + 1) \). A maximum of four snapshots are temporarily stored in the OBU before recording the start and end snapshots of a virtual segment. The various stages of R² protocol are presented using a flowchart in Figure 2. When OBU starts recording snapshots, the first, second, and third speed observations and other snapshot parameters such as location and time are temporarily stored in \( S_1 \), \( S_2 \), and \( S_f \). A linear regression line is fitted to the speed observations of \( S_1 \), \( S_2 \), and \( S_f \), and the R² of the fitted line is computed. Based on the R² value two situations arise:

- If the R² is smaller than the acceptable threshold then \( S_1 \) is permanently recorded in the OBU to be used as the starting snapshot of the virtual segment, then \( S_2 \) is assigned to \( S_1 \), and \( S_f \) is assigned to \( S_2 \) and then a new virtual segment is started.

- If the R² is greater than the acceptable threshold then a new snapshot is temporarily recorded and assigned to \( S_{f+1} \).

In the next step, a linear regression line is fitted to \( S_1 \), \( S_2 \), \( S_f \), and \( S_{f+1} \), and R² of the regression line is calculated. If the R² value is greater than the desired threshold, \( S_f \) is replaced with \( S_{f+1} \) and as vehicle continues traveling, the next available speed observation is assigned to \( S_{f+1} \). The regression procedure is repeated and as long as the R² value is above the desired threshold, the value of \( S_{f+1} \) is assigned to \( S_f \). When the new \( S_{f+1} \) results in an R² value lower than the threshold, the speed and other parameters of snapshot \( S_1 \) are permanently recorded in the OBU, and \( S_f \) is assigned to \( S_1 \) (i.e., a new virtual segment is started). In other words, the two snapshots \( S_1 \) and \( S_f \) represent one virtual segment of the time-speed plot. When a vehicle enters the coverage range of an RSE, all the recorded snapshots are transferred to the RSE and purged from OBU memory.
* s = attributes of a probe vehicle, including vehicle ID, current speed, position, simulation time and link ID.
1 $S_1, S_2, S_f$: temporary variables for storing snapshots
2 $R^2$: threshold of acceptable coefficient of determination for snapshot segments, OBU: on-board unit

FIGURE 2 $R^2$ Protocol
Travel Time Estimation Using Reconstructed Time-Speed Plot

Snapshots include information such as vehicle speed, vehicle location, and the time when a snapshot was recorded. The time-speed plot is generated by interpolating the data obtained from recorded snapshots. Interpolation can be used to increase the accuracy of travel time estimation. Interpolation methods such as linear, cubic spline (29), and cubic Hermite are used (30-31) in this paper. Through interpolation the earliest and latest times at which a vehicle is on a link are obtained. This section explains a method to generate link travel time for a vehicle using the time-speed plot. From the time-speed plot one can discretize the time axis into 1 second intervals. Figure 3 illustrates the \( n \) time points during which \( \text{veh}_k \) is on link \( l \). For \( \text{veh}_k \), position \((p)\) at various times \((t)\) can be derived using speed \((v)\). Travel time of \( \text{veh}_k \) on link \( l \) is calculated according to equation (1):

\[
LTT_k = \frac{p_1^{\text{veh}_k} - l_0}{v_1^{\text{veh}_k}} + (t_n^{\text{veh}_k} - t_1^{\text{veh}_k}) + \frac{t_e - p_n^{\text{veh}_k}}{v_n^{\text{veh}_k}}
\]

where,

- \( LTT_k \): link travel time of \( \text{veh}_k \)
- \( l_0 \): coordinates of beginning of the link
- \( p_1^{\text{veh}_k} \): position of first instance of \( \text{veh}_k \) on link \( l \)
- \( v_1^{\text{veh}_k} \): speed of \( \text{veh}_k \) in first instance on link \( l \)
- \( t_n^{\text{veh}_k} \): last instance of \( \text{veh}_k \) on link \( l \)
- \( t_1^{\text{veh}_k} \): first instance of \( \text{veh}_k \) on link \( l \)
- \( l_e \): coordinates of end of the link
- \( p_n^{\text{veh}_k} \): position of first instance of \( \text{veh}_k \) on link \( l \)
- \( v_n^{\text{veh}_k} \): speed of \( \text{veh}_k \) in first instance on link \( l \)

Assuming \( m \) vehicles travel through link \( l \) during \( i^{th} \) travel time estimation interval, link travel time during the \( i^{th} \) time interval is estimated as:

\[
ELT_{T_{l,i}} = \frac{\sum_{k=1}^{m} LTT_{k,i}}{m}
\]

where, \( ELT_{T_{l,i}} \) = estimated travel time for link \( l \) during the \( i^{th} \) time interval
\( m \): number of vehicles traveling through the link \( l \) in the \( i^{th} \) time interval

\( LTT_{i,h} \): travel time of vehicle \( h \) on link \( l \) in the \( i^{th} \) time interval

**CASE STUDY APPLICATIONS**

The simulation test bed of downtown Boise, Idaho, network developed in a previous study (6) was used in this study to compare the performance of different protocols. The VISSIM Component Object Model (COM) feature was used to obtain vehicles’ speed and position data in 1-second intervals. The road network consisted of 152 links and the average link length was 389 feet (0.07 mile). The network demand during the simulation period was 22,618 vehicles. It was assumed that RSEs are installed at every intersection and connected vehicle market penetration was 100%. The Fixed 2 sec (F2), Fixed 4 sec (F4), SAE J2735 and \( R^2 \) probe data collection protocols were implemented in VISSIM and snapshots were recorded. Linear regression, cubic spline interpolation, and piecewise cubic Hermite interpolation were used to construct time-speed plots. The link travel times were estimated every two minutes. Ground truth travel times were obtained from simulation every one second. Mean absolute relative error (MARE), was computed for the estimated travel times as follows:

\[
MARE_l = \frac{1}{N} \sum_i \left| \frac{E_{TT_{l,i}} - E_{TT_{l,i}}}{T_{TT_{l,i}}} \right|
\]

where,

- \( MARE_l \) = mean absolute relative for link \( l \)
- \( N \) = total number of travel time estimation intervals
- \( E_{TT_{l,i}} \) = estimated travel time for link \( l \) in interval \( i \)
- \( T_{TT_{l,i}} \) = ground truth travel time of link \( l \) in time interval \( i \)

Figure 4 shows an example of snapshots collected by each protocol and the ground truth data for a link in the Boise network. As previously mentioned, F2 and F4 protocols collect redundant snapshots when a vehicle is stopped, on the other hand, the snapshots collected using the SAE protocol do not necessarily provide adequate information for instances in which a vehicle is stopped or moves slower than 10 mph. In Figure 4, snapshots collected using \( R^2 \) protocol seem to better represent the time-speed plot of a vehicle by capturing the changes in vehicle speed and by collecting adequate samples in slow moving or stopped conditions. This hypothesis is later evaluated by comparing link travel time estimate obtained from \( R^2 \) protocol with travel time estimates obtained from F2, F4, and SAE protocols.

The travel time on each link in the network was estimated for 37 time intervals. Table 1 presents the mean and variance of the MARE measure for the entire network. The first three rows in Table 1 provide the results for the three interpolation methods for the F2, F4, and SAE protocols. The results of \( R^2 \) protocol are presented later. The fourth row in Table 1 provides travel times estimated using the ‘average speed method’. In the average speed method, link travel time is estimated by dividing the length of the link by average speed of all uploaded snapshots of vehicles travelling on that link (6) during that time interval.
Based on the values shown in Table 1, the piecewise cubic Hermite method outperformed the other methods. Statistical tests, t-test for mean and F-test for variance, revealed that the differences were statistically significant at $p = 0.05$. Cubic spline interpolation performed better than the linear interpolation method for all protocols.

The number of snapshots collected by F2, F4, and SAE was not the same. The F2, F4, and SAE protocols generated 1,424,631, 716,957, and 309,398 snapshots, respectively. For a fair comparison between the $R^2$ protocol and the other protocols, the number of collected snapshots in the $R^2$ protocol should be equal to or smaller than the number of snapshots collected by the other protocol. Thus, three additional variations of the $R^2$ protocol were generated to match the number of snapshots collected in the other protocols. The thresholds for primary $R^2$ protocol and the three variations of the protocol are presented in Table 2. In the primary $R^2$ protocol, referred to as $R^2$-Group 1 in Table 2, snapshots are collected such that $R^2$ is always 0.95. In other variations of $R^2$ protocol, referred to as $R^2$-Group 2 to $R^2$-Group 4, the thresholds were obtained by varying the threshold such that the number of collected

### TABLE 1 Mean and Variance of MARE for the Boise network

<table>
<thead>
<tr>
<th>Method</th>
<th>F2</th>
<th>F4</th>
<th>SAE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Variance</td>
<td>Mean</td>
</tr>
<tr>
<td>Linear</td>
<td>0.0368</td>
<td>0.0006</td>
<td>0.0995</td>
</tr>
<tr>
<td>Cubic Spline</td>
<td>0.0227</td>
<td>0.0002</td>
<td>0.0703</td>
</tr>
<tr>
<td>Piecewise Cubic Hermite</td>
<td>0.0160</td>
<td>0.0001</td>
<td>0.0439</td>
</tr>
<tr>
<td>Average Speed</td>
<td>0.0872</td>
<td>0.0040</td>
<td>0.1297</td>
</tr>
</tbody>
</table>
TABLE 2  $R^2$ Protocol Thresholds

<table>
<thead>
<tr>
<th>R2 protocol variations</th>
<th>Min(speed) $\geq$ 15 ft/s</th>
<th>Max(speed) $\geq$ 15 ft/s</th>
<th>Max(speed) $\leq$ 15 ft/s</th>
<th>Mean(speed) $\leq$ 7 ft/s</th>
<th>Max(speed) $\leq$ 15 ft/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$-Group 1</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>$R^2$-Group 2</td>
<td>0.95</td>
<td>0.9</td>
<td>0.9</td>
<td>0.5</td>
<td>0.9</td>
</tr>
<tr>
<td>$R^2$-Group 3</td>
<td>0.8</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$R^2$-Group 4</td>
<td>0.6</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Snapshots would be equal to or smaller than the snapshots collected in the F2, F4, and SAE protocols. Four $R^2$ threshold values were defined for each $R^2$-Group. $R^2$ threshold values are reported in columns 2 to 5 of Table 2. Each column in Table 2 represents a speed category; for example column 2 represent the instances where speed of vehicle in snapshots segment is greater than or equal to 15 ft per second (approximately 10 mph). Exhaustive search method was used to determine the threshold values for each $R^2$-Group. Higher $R^2$ threshold values were assigned to higher speed groups (e.g., Groups 1 and 2) since errors in recording vehicle speed in these groups would have a higher effect on estimated travel times.

The travel time estimation error results for the four $R^2$ protocol groups are shown in Table 3. To facilitate comparison, results of F2, F4, and SAE protocols are also included in the table. The least travel time estimation error was obtained when $R^2$ protocol was used for snapshot generation and cubic Hermite interpolation method was used for time-speed plot generation. T-test was used to compare the mean of MARE obtained from cubic Hermite interpolation between: 1) F2 protocol and $R^2$-Group 2, 2) F4 protocol and $R^2$-Group 3, and 3) SAE protocol and $R^2$-Group 4 protocol. The results indicate that with $p = 0.00$, combination of $R^2$ protocol and cubic Hermite interpolation outperformed the F2, F4, and SAE protocol, for same or fewer number of snapshots. These results also confirmed the earlier hypothesis that the $R^2$ protocol probe data best describe the time-speed plot of individual vehicles.

Furthermore, sensitivity analysis was conducted with respect to the market penetration rates and travel time estimates obtained from SAE J2735 and $R^2$ protocols under 20%, 50%, 80%, and 100% market penetration rates are compared in Table 4. As shown in column 3 of Table 4, $R^2$ protocol collected fewer snapshots in comparison with SAE J2735. In terms of travel

TABLE 3  Comparison of $R^2$ Protocol with F2, F4, and SAE Protocols using Boise, ID Simulation test-bed

<table>
<thead>
<tr>
<th>Probe Data Collection Protocol</th>
<th>Number of Recorded Snapshots</th>
<th>Mean of MARE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Linear</td>
</tr>
<tr>
<td>$R^2$-Group 1</td>
<td>1,076,600</td>
<td>0.069</td>
</tr>
<tr>
<td>$R^2$-Group 2</td>
<td>995,709</td>
<td>0.070</td>
</tr>
<tr>
<td>F2</td>
<td>1,424,957</td>
<td>0.037</td>
</tr>
<tr>
<td>$R^2$-Group 3</td>
<td>530,418</td>
<td>0.183</td>
</tr>
<tr>
<td>F4</td>
<td>716,957</td>
<td>0.010</td>
</tr>
<tr>
<td>$R^2$-Group 4</td>
<td>295,823</td>
<td>0.252</td>
</tr>
<tr>
<td>SAE</td>
<td>309,398</td>
<td>0.268</td>
</tr>
</tbody>
</table>
**TABLE 4** Effect of Market Penetration Rate on Travel Time Estimation Error

<table>
<thead>
<tr>
<th>Market Penetration Rate</th>
<th>Probe Data Collection Protocol</th>
<th>Number of Recorded Snapshots</th>
<th>Mean of MARE for travel time estimation method:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Linear</td>
</tr>
<tr>
<td>100%</td>
<td>R² *</td>
<td>295,823</td>
<td>0.252</td>
</tr>
<tr>
<td></td>
<td>SAE</td>
<td>309,398</td>
<td>0.268</td>
</tr>
<tr>
<td>80%</td>
<td>R² *</td>
<td>235,716</td>
<td>0.263</td>
</tr>
<tr>
<td></td>
<td>SAE</td>
<td>246,517</td>
<td>0.274</td>
</tr>
<tr>
<td>50%</td>
<td>R² *</td>
<td>147,221</td>
<td>0.269</td>
</tr>
<tr>
<td></td>
<td>SAE</td>
<td>153,989</td>
<td>0.282</td>
</tr>
<tr>
<td>20%</td>
<td>R² *</td>
<td>57,923</td>
<td>0.274</td>
</tr>
<tr>
<td></td>
<td>SAE</td>
<td>60,815</td>
<td>0.292</td>
</tr>
</tbody>
</table>

*R² protocol with threshold values same as R²-Group 4 in Table 2 is used to collect probe data*

Time estimation error, combination of R² protocol probe data and Cubic Hermite travel time estimation method produced the smallest travel time estimation error and consistently outperformed SAE J2735 probe data for 20%, 50%, 80%, and 100% market penetration rates. The protocol evaluations were conducted using a computer cluster. The computation time for the entire simulation period for all vehicles was 11.6 seconds for the SAE protocol and 19.8 seconds for the R² protocol.

In a second case study, real world data from third major trial of connected vehicles at the U.S. DOT Michigan test-bed was used to evaluate the performance of R² protocol. The data is part of National Center for Atmospheric Research (NCAR) connected vehicle dataset; the dataset and detailed description of data were obtained from the FHWA research data exchange website (32). Several vehicles equipped with OBU units drove through the test-bed for several days. Snapshots were collected according to the SAE J2735 protocol, and vehicle trajectories were collected in 1-second intervals. Trajectory data and snapshots include timestamp, speed, latitude and longitude of vehicles. Data collected by three probe vehicles identified as “B042”, “C695” and “C832” in the NCAR dataset were used in this case study. Test vehicles traversed links 20, 22, 23, and 25 of test-bed multiple times between 5:00 pm and 6:00 pm on March 29, 2010. For simplicity, the three test vehicles are referred to vehicle 1, 2, and 3, and the links are referred to as link 1, 2, 3, and 4. The average link length was 1.3 miles. Figure 5(a) shows the location of links 1, 2, 3 and 4. Figures 5(b) to 5(d) illustrate the locations of snapshots collected by vehicles 1, 2, and 3 according to the SAE J2735 protocol. Although the test vehicles traversed the links multiple times, Figure 5 only shows the snapshots collected during one passage on the link.

The R² protocol was applied to the trajectory data and snapshots were generated. The locations of R² snapshots for vehicle 1, 2, and 3, for one link passage, are also shown in Figure 5 (b) to Figure 5 (d). The generated snapshots were used to estimate travel time of test vehicles on the study links. Travel times were also estimated using the SAE J2735 snapshots collected from the field. Ground truth travel time of test vehicles was calculated from 1-second trajectory data. The mean absolute relative error (MARE) of test vehicles’ travel time estimated using R² and SAE J2735 snapshots is reported in Table 5. The results indicated that the R² protocol with cubic Hermite method produced the smallest travel time estimation error. Also, the total number of snapshots for all four links collected by the R² protocol,
FIGURE 5 Location of $R^2$ Snapshots versus SAE J2735 Snapshots for vehicles 1, 2, and 3 in Michigan test-bed

TABLE 5 Comparison of $R^2$ Protocol with SAE Protocols using U.S. DOT Michigan test-bed data

<table>
<thead>
<tr>
<th>Link Number</th>
<th>Probe Data Collection Protocol</th>
<th>Number of Recorded Snapshots</th>
<th>MARE Linear</th>
<th>Cubic Spline</th>
<th>Cubic Hermite</th>
<th>Average Speed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SAE</td>
<td>85</td>
<td>0.067</td>
<td>0.154</td>
<td>0.021</td>
<td>0.672</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>153</td>
<td>0.024</td>
<td>0.034</td>
<td>0.012</td>
<td>0.460</td>
</tr>
<tr>
<td>2</td>
<td>SAE</td>
<td>129</td>
<td>0.191</td>
<td>0.079</td>
<td>0.092</td>
<td>0.667</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>73</td>
<td>0.237</td>
<td>0.081</td>
<td>0.023</td>
<td>0.832</td>
</tr>
<tr>
<td>3</td>
<td>SAE</td>
<td>100</td>
<td>0.048</td>
<td>0.165</td>
<td>0.066</td>
<td>0.568</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>69</td>
<td>0.024</td>
<td>0.040</td>
<td>0.026</td>
<td>0.619</td>
</tr>
<tr>
<td>4</td>
<td>SAE</td>
<td>91</td>
<td>0.204</td>
<td>0.048</td>
<td>0.084</td>
<td>1.191</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>62</td>
<td>0.285</td>
<td>0.117</td>
<td>0.030</td>
<td>1.213</td>
</tr>
<tr>
<td>All Links</td>
<td>SAE</td>
<td>405</td>
<td>0.128</td>
<td>0.111</td>
<td>0.066</td>
<td>0.775</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>357</td>
<td>0.142</td>
<td>0.068</td>
<td>0.023</td>
<td>0.781</td>
</tr>
</tbody>
</table>
shown in the third column and last two rows of Table 5, were 11% fewer than the SAE J2735 snapshots. As shown in Figures 5(b) and 5(d) comparing to the SAE J2735 protocol, fewer snapshots were recorded by the R^{2} protocol on mid-block sections of links 3 and 4 where vehicle speed is not expected to vary significantly, whereas more snapshots were recorded near intersections where significant speed changes occurred.

CONCLUSION

Probe traffic data collection is one of the promising features of connected vehicle technology. Vehicle OBU record snapshots that include information such as vehicle speed, position, according to a data collection protocol. A novel probe data collection protocol called the R^{2} protocol was presented in this paper. In the existing protocols, snapshots are recorded in fixed time intervals or in a predefined range of intervals. In the R^{2} protocol, snapshots are collected only when a significant change occurs in the vehicle speed. The performance of the R^{2} protocol was evaluated using a simulated case study of downtown Boise, and using real-world data collected in the U.S. DOT connected vehicle test bed in Michigan.

The proposed R^{2} protocol was found to perform better than three existing protocols, the Fixed 2 second, Fixed 4 second, and SAE J2735 protocols. The better performance is achieved by avoiding collecting redundant snapshots during stop events, and capturing the critical snapshots when significant speed changes occur. Not only did the R^{2} protocol outperform the three protocols in terms of the error measure but it also had required fewer snapshots to achieve that. The number of snapshots recorded by the R^{2} protocol in downtown Boise was 30%, 26% and 4% lower than the Fixed 2 second, Fixed 4 second, and SAE J2735 protocol, respectively. Application of R^{2} protocol to the Michigan test-bed data also confirmed that R^{2} protocol outperformed the SAE J2735 probe data in terms of number of collected snapshots and travel time estimation error. The promising results found in this research encourage future research in applying the R^{2} protocol to additional case studies to further validate the protocol performance. The performance of R^{2} protocol for other traffic management and traveler information applications also needs to be investigated in future research.

REFERENCES


