Traffic State Estimation using DSRC-Enabled Probe Vehicles

Sarthak Chaturvedi, Abirami Krishna Ashok, Dr. Bhargava Rama Chilukuri

Introduction

- Traffic estimation refers to the observation of three fundamental variables - flow, density and speed - also known as traffic state.
- Characteristics of traffic observation method can be broadly explained under three heads:
    - If all the fundamental variables could be acquired simultaneously
    - If wide spatial and temporal area could be covered
    - If rich and accurate data is acquired for spatial-temporal area covered
- Location based sensors like loop detectors and video cameras acquire the traffic state information at only the installed locations
- Probe vehicles have onboard data collection devices that record traffic information while the vehicle is in traffic flow.
    - Commonly used probe vehicles are equipped with GPS device to identify position and speed of the vehicle.
    - However, GPS enabled probe vehicle could not acquire flow and density information

Literature Review

- ➔ Seo et al. (2015a) utilized GPS equipped probe vehicles that can measure spacing with the leader vehicle for state estimation
  - The traffic state estimation was carried out by discretizing time-space region
  - State variables are estimated in each region by using Edie's generalized definitions
  - The result showed that the proposed method could accurately estimate the 5 min and hourly traffic volumes
- ➔ Ding, F.(2019) explored the use of WiFi signal data, for travel time estimation.
  - WiFi detectors were installed along a an highway to record signal data from WiFi activated devices inside vehicles.
  - Mining and filtering algorithms were used to extract necessary data to estimate state variables - speed and travel time.
  - Comparisons of the results with traffic speed data obtained from loop detectors were found to be consistent
  - A highly correlated relationship was also observed between estimated traffic volume and the loop detector data
- ➔ Patra, S. et al. (2019) worked on development of a low cost WiFi MAC Scanner (WMS) for traffic application.
  - The range of WMS was estimated to be approximately 80m.
  - Variation of received signal strength indicator (RSSI) with increase in distance (D) was found to be logarithmic in nature.
  - The developed WMS had good detection performance and higher matching rates for traffic applications

Gaps in Literature

- While the literature shows that state estimation with probe vehicles is promising, there are several drawbacks of the existing methods in terms of sensors, data, communication, and methodology.
- The sensors used by the probe vehicles are often expensive and cumbersome to install if it didn’t come pre-installed in the vehicle.
- In terms of the data collection, these sensors often detect only the vehicles in their immediate vicinity.
- Moreover, the communication technology used by the studies is expensive since each vehicle communicates the data to the server using General Packet Radio Service (GPRS) or similar technology.
- Finally, the limited sample data used by these methodologies make the state estimates unreliable and can result in wide variations across consecutive time periods.
- To overcome the limitations of existing sensors, an estimation method based on mobile sensors and vehicle to infrastructure (V2I) communication system is proposed
  - Probe vehicles capable of observing surrounding vehicles by the Media Access Control (MAC) identifier are used
  - WiFi sensors embedded in these vehicles can acquire data of a wide spatio-temporal region
  - Limited instrumentation required and cost effective

Methodology

- ➔ This study aims to develop a probe vehicle-based traffic state estimation method.
- ➔ In this study, municipal corporation buses are used as probe vehicles.
- ➔ These buses are instrumented with on-board units (OBU) with integrated WiFi sensors to acquire information of any WiFi devices in its proximity
- ➔ To regularly collect the data from probe vehicles, roadside units(RSUs) are installed in key strategic locations along the length of the corridor.
- ➔ OBUs use Dedicated short-range communications (DSRC) technology to transmit the data to the roadside units(RSUs).
- ➔ The data packets from OBUs are transmitted to the RSUs whenever the former comes in the vicinity of the latter.
- ➔ These RSUs transfer the collected data to remote servers for quantitative state estimation.

Data Collection

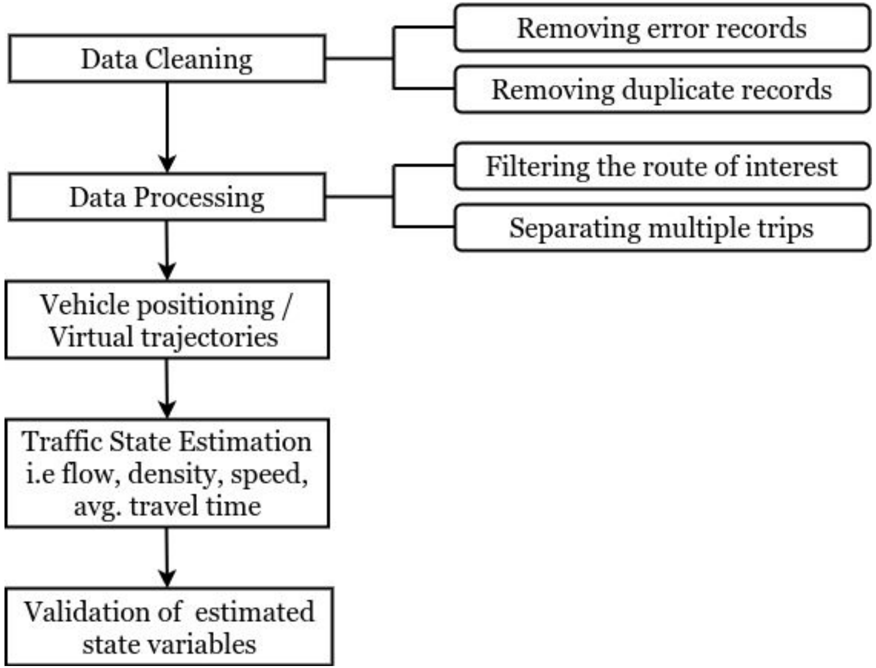
Table I shows a sample data set recorded by the OBU. This data broadly consist of following parameters:

- ➔ Source ID: Helps in identifying the probe vehicle from which data is obtained.
- ➔ MAC-ID, Signal Strength: Collects the unique identification of WiFi devices and corresponding signal strength. This signal strength is used to calculate space headway between the probe vehicle and other Wi-Fi devices.
- ➔ GPS coordinates (lat, long): Helps in identifying the exact location of probe on the ground.
- ➔ UTC Date and Time stamp: Identifies time corresponding to the given location of the probe.

TABLE I: WiFi signal data attributes.

Attributes Description	Sample Data	Remarks
Source ID	1	OBU unique ID
MAC address	00:0a:95:9d:86:17	
Signal Strength	-50	Range from -1 to -99
GPS Location	8.51, 76.56	latitude, longitude
UTC Date	281107	HHMMSS
UTC Time	061241	DDMMYY

Schematic of the proposed traffic state estimation system.



Data Cleaning

- Removing duplicate records
  - Occasionally, duplicates MAC records are sent by devices in the WiFi channel.
  - Only one of the records (reliable) with a unique MAC address is retained, and others are eliminated.
  - For records with identical MAC address, the MAC address recorded at the lowest signal strength was used for analysis (i.e., readings that place the Wi-Fi device at the farthest distance from the probe vehicle).

- Removing error records
  - Occasionally, faults may occur in communication systems leading to incorrect formatting of records such as an unreasonable timestamp, inconsistent MAC recording, etc.
  - These errors are identified and are removed while processing raw data.

V2I,31,00,02,88,000022,A,12.84910333,80.06425833,161020,16:47
:2D:B6:FD:F9,-81,16:47:2D:B6:FD:F9,-53,1668,abcd

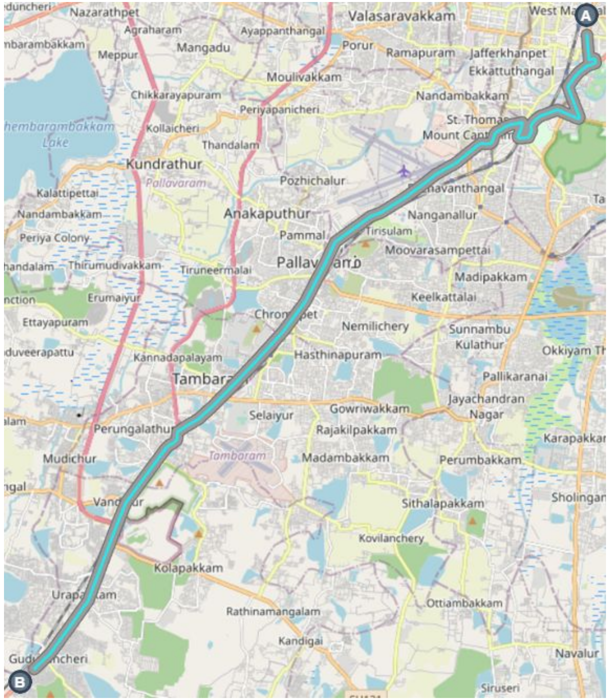
V2I,31,00,02,134,180855,A,13.034225,80.23035,151020,
F8:E7:1E:4B:68:23,-67,3E:53:DA:32:55:C8,-57,-26 10:31:03,(not associated) ,3E:53:DA:32:55:C8,-57,61,abcd

<u>Sample error records</u>

Data Processing

Once the input data is cleaned, the next objective is to process it to make it usable for the analysis. There are several steps involved in the pre-processing of the raw data.

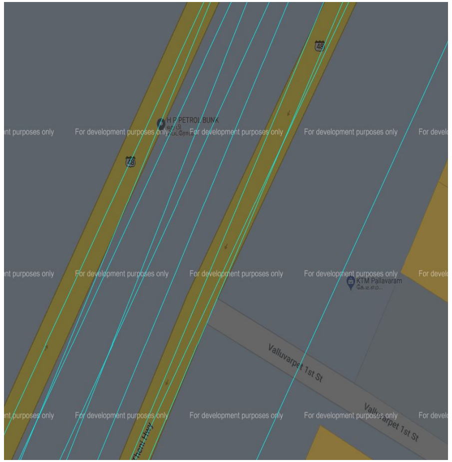
- ➔ Filtering the route of interest
  - ◆ The city buses used as probe vehicles may not always travel along the route of our interest.
  - ◆ To overcome this challenge, a buffer region of 100m is created along the centerline of the route of interest.
  - ◆ Then, the GPS data is intersected with the buffer region to strip any non-essential recording.



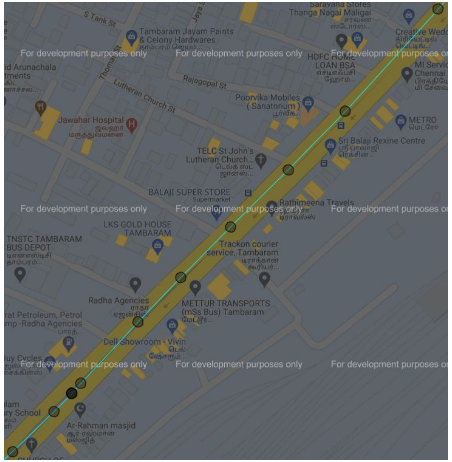
Filtering of probe route

Data Processing

➔ Separating multiple trips
- Another hurdle is to classify data into separate trips based on the direction of movement of the probe.
- A mining algorithm is devised that contains GPS coordinates of landmarks arranged in order along the probe's direction of movement such that whenever the probe hits the first landmark, recording of data starts until the next landmark is reached.



Initially



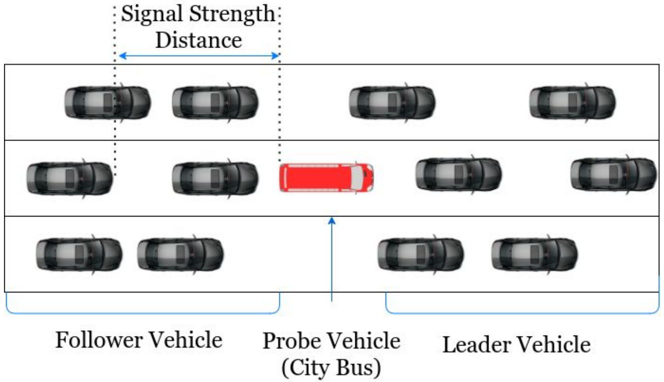
Results post separation of trips

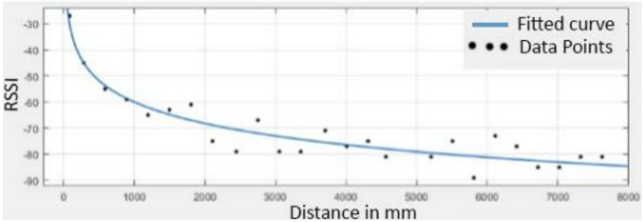
Distribution of WiFi Devices around probe

- ➔ Only the MAC IDs which are recorded by OBU multiple times are used for distribution.
- ➔ Half of the these MAC IDs are designated as leader and the other half as follower vehicles at random to avoid any bias.
- ➔ The spacing between the probe and the surrounding vehicles is estimated based on the signal strength of device.
- ➔ Based on work of Patra, et al (2019) relation between signal strength and distance of transmitting device was established. Reported studies showed a logarithmic decrease in RSSI value with an increase in the distance.

$$D = 10^{(A - RSSI)/B}$$

  - ◆ D = Distance of transmitting device (cm)
  - ◆ RSSI = Received Signal Strength Indicator (dBm)
  - ◆ A,B = Calibration parameters
- ➔ Value of calibration parameter for WiFi sensors to capture signals from a maximum distance of 75 m are:
  - ◆ A = 37.19; B = 33.59

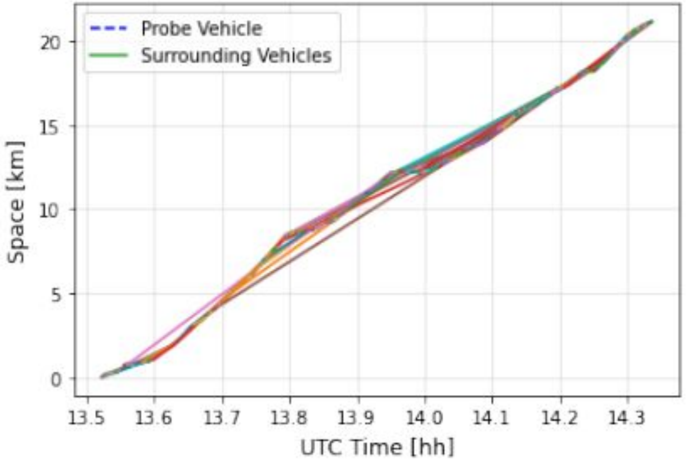




Scatter plot of distance with RSSI
Source: Patra et al. (2019)

Virtual Trajectories

- ➔ Post the distribution of vehicles the space-time (ST) trajectories are generated.
- ➔ Since GPS coordinates of probe are available at every 6s interval and hence, generating it's ST plot is easy.
- ➔ In contrast, the surrounding vehicles have very few data points to form an accurate trajectory.
- ➔ Surrounding vehicles are plotted with the probe vehicle trajectory as the reference.
- ➔ Data point corresponding to a MAC ID is placed above or below probe's trajectory based on its classification as leader (or follower)
- ➔ The magnitude by which trajectory is above or below the probe depends on its signal distance of transmitting device.

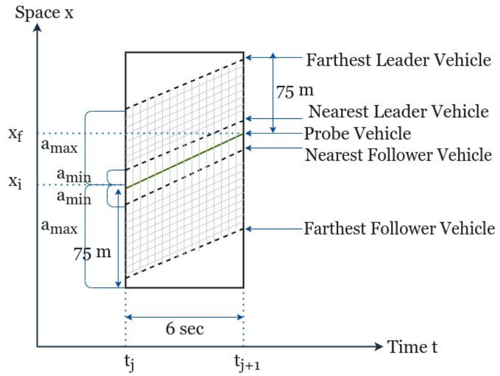


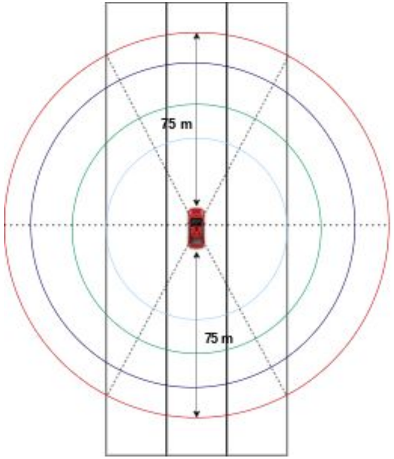
Space-TIme diagram

Traffic State Estimation

The proposed method estimates traffic state variables at a pre-stated timespace resolution along the probe trajectory.

- ➔ ST diagram is divided into multiple smaller discrete time-space regions.
- ➔ Discrete time-space region study recorded data at every 6-sec interval and in a region 75 m behind the probe's initial position and at 75 m ahead of probe's final position.
- ➔ This limit of 75 m is after the fact that the OBU can record farthest WiFi signal at this distance.
- ➔ The frequency at which data is available might not be 6 sec for all the recordings.
- ➔ To keep our study and time-space entity consistent interpolation of data points between two known data points.





Modified Edie's Definition

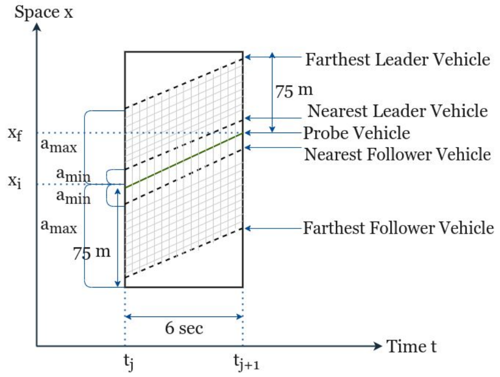
The traffic state variables - namely, the flow q(A), density k(A) and speed v(A) as defined by Seo (2015a) utilizing Edie's generalized definitions are used in this study.

- ➔ N(A) represents all vehicles in the space-time region A.
- ➔ $\sum_{n \varepsilon N(A)} d(A)$ is the distance travelled by all the vehicles in region A (veh-km).
- ➔ $\sum_{n \varepsilon N(A)} t(A)$ is the time spent by all the vehicles in region A (veh-hr).
- ➔ $\sum_{n \varepsilon N(A)} |a(A)|$ is the sum of area between the trajectory of a surrounding vehicle and the probe in the region A.

$$\hat{q}(A) = \frac{\sum_{n \varepsilon N(A)} d(A)}{\sum_{n \varepsilon N(A)} |a(A)|}$$

$$\hat{k}(A) = \frac{\sum_{n \varepsilon N(A)} t(A)}{\sum_{n \varepsilon N(A)} |a(A)|}$$

$$\hat{v}(A) = \frac{\sum_{n \varepsilon N(A)} d(A)}{\sum_{n \varepsilon N(A)} t(A)}$$



Time-space region

Proposed State Estimation Method

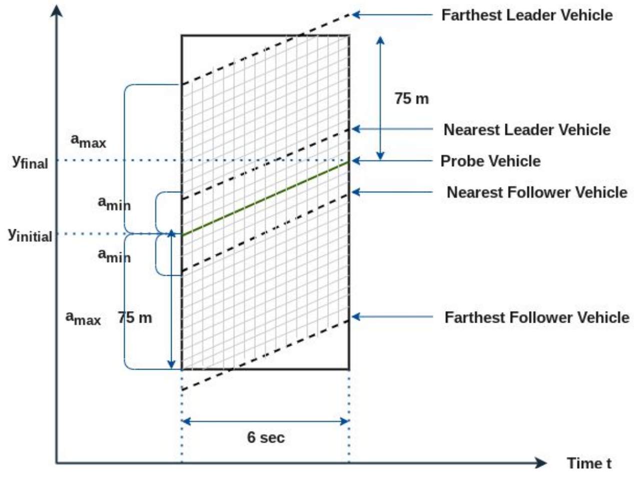
In this study we propose two methods to calculate the traffic state:

- ➔ Method 1: Farthest leader and follower configuration
- ➔ Method 2: Nearest leader and follower configuration
- ➔ The area of interest for the Method 1 and 2 as defined as:
- ➔

$$A^M = a_l^M + a_f^M$$

$$A^m = a_l^m + a_f^m$$

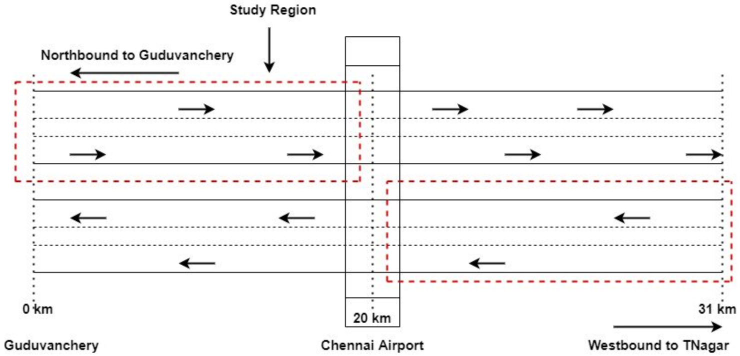
- ◆ A^M (or A^m) is the area between the farthest (or nearest) leader and farthest (or nearest) follower.
- ◆ a_l^M (or a_l^m) is the area between the farthest (or nearest) leader and probe.
- ◆ a_f^M (or a_f^m) is the area between the farthest (or nearest) follower and probe.



Time-space region

Field Experiment

- ➔ Conducted on October 16, 2020, from 00:20 to 14:20 UTC
- ➔ Testbed is a bidirectional urban highway leading from Guduvanchery to Chennai International Airport.
- ➔ Probe vehicle, a municipal city bus, started from the T-Nagar depot and shuttles to and from Guduvanchery via Airport during each trip.
- ➔ Three RSUs are installed at strategic positions along this highway to transmit real-time data.
- ➔ Probe vehicle recorded data at an interval of 6 s, and on average, 389 data packets were recorded per trip from Guduvanchery to Airport.



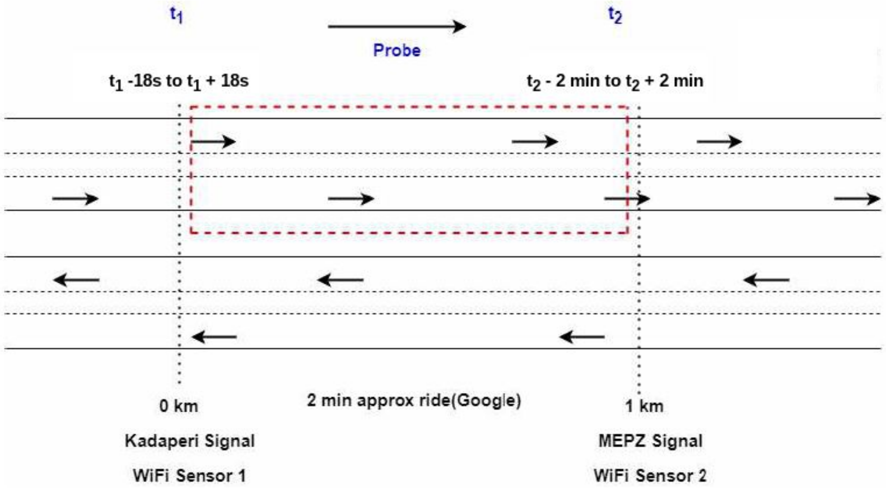
Testbed

Attributes	Trip 1	Trip 2	Trip 3	Trip 4	Trip 5
Total MAC detected	1281	2641	2160	2040	2727
Unique MAC detected	940	2051	1678	1607	2143
MAC detected multiple times	114	178	174	195	215

Descriptive analysis of recorded data

Validation

- For comparing traveltime, independent WiFi sensors are installed at Kadaperi and MEPZ signal areas.
- The corridor between these locations is a driving stretch of approximately 1 km.
- MAC Ids recorded at sensor 1 within +/- 18 seconds of the probe's entry time at the location were considered for validation.
- Similarly, the time these MAC IDs reidentified at sensors 2 within +/- 2 minutes of probe's exit time from the location were considered.
- The reason for providing a larger recording window at the downstream location is to allow capturing of the fastest and slowest vehicle in the traffic stream.

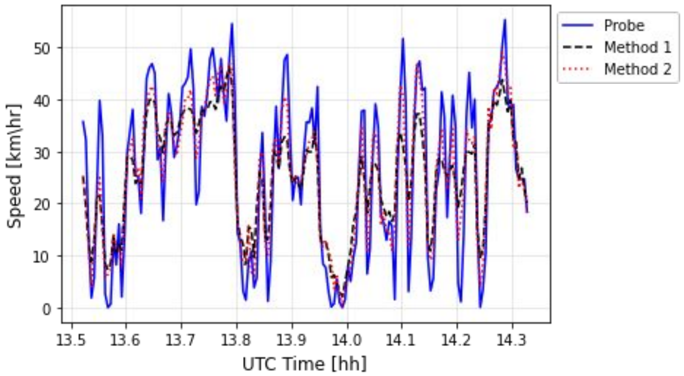


Validation testbed

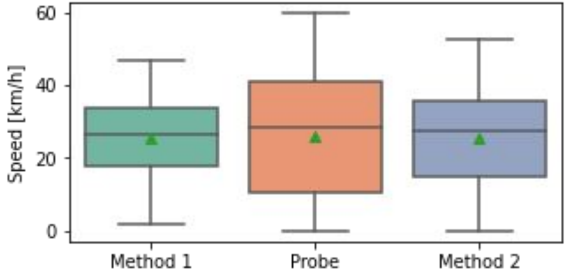
Results

The probe shuttled 5 times from the Guduvanchery to the airport. Data recorded during trip 5 is used to estimate traffic state.

- ➔ To evaluate the performance of estimates made through Method 1 and 2, a comparison of average speed estimated by these methods with the average speed of probe is done.
- ➔ Results suggest that the speed estimates using Method 2 better represents the probe's speed.
  - ◆ Method 1 represents the average speed of larger number of vehicles in the time-space entity.
  - ◆ In contrast, Method 2 estimates the average speed of the probe and its immediate leader and follower
- ➔ A statistical analysis showed that the mean speed for both methods is approximately +/- 3 km/hr of probe's speed.
- ➔ However, method 2 was able to capture the higher and lower extremes of probe's speed more efficiently than Method 1.



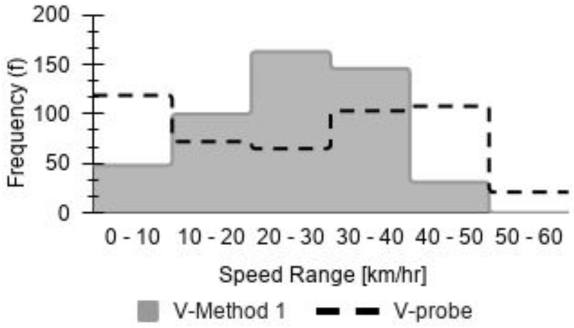
Comparison of estimated speed



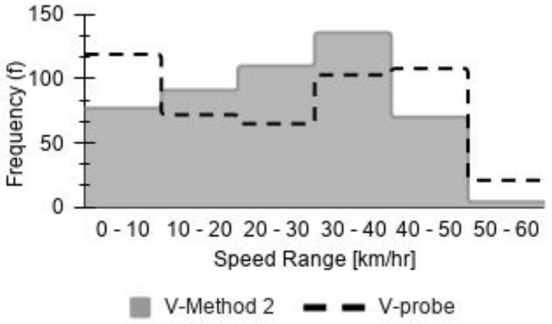
Results

A comparison of the speed distribution in the probe and the estimated speed using Method 1 and 2 was studied.

- ➔ It could be observed that Method 2 reflected small variations in speed of probe (or traffic) more effectively than method 1.
- ➔ Estimated speed using Method 2 depicts closer co-relation in speed distribution with the probe when compared with estimated speed using Method 1.
- ➔ Upon conducting further statistical analysis, the strength of the association between estimated speed using Method 2 and average speed of probe vehicle across the entire corridor, measured using Pearson Correlation coefficient is found to be **0.93**.
- ➔ For Method 1, this statistic is found to be **0.875**



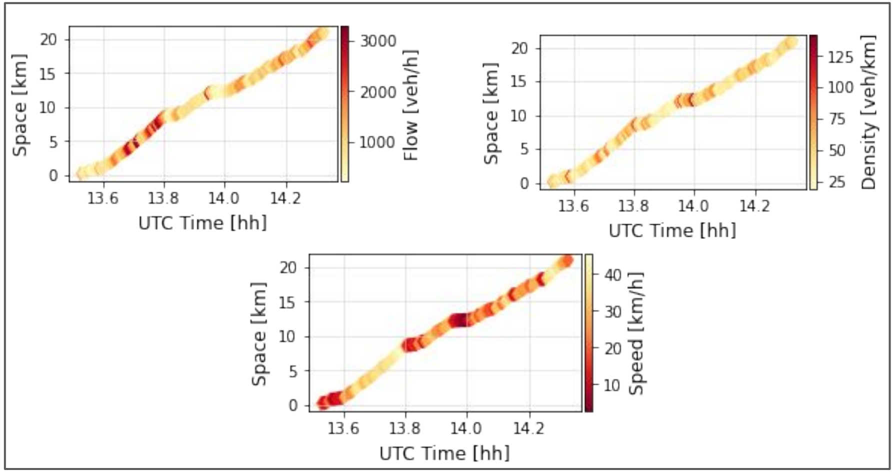
Speed distribution in Method 1

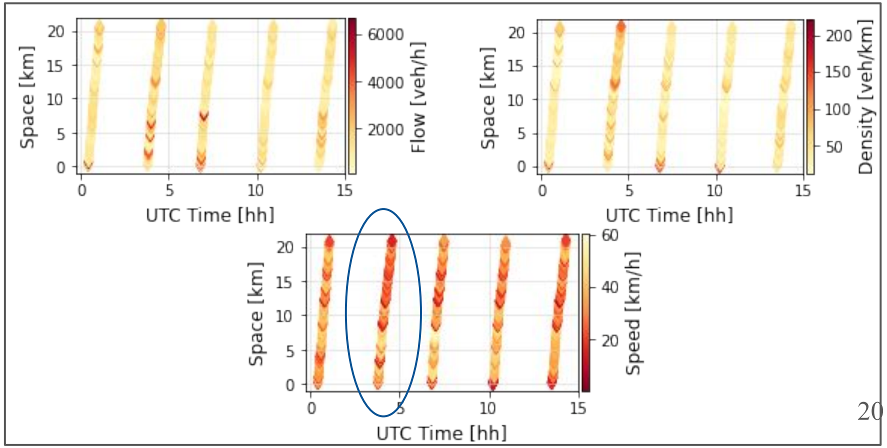


Speed distribution in Method 2

Results

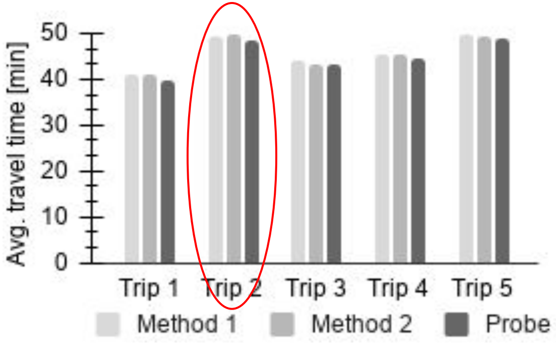
- ➔ Although speed estimation using Method 2 showed a better co-relation with the probe's speed, however, the estimated flow and density using this method resulted in unreasonably high values.
- ➔ Estimated flow and density using Method 1 showed values within the practical range.
- ➔ The traffic state estimated using Method 1 for the all the trips done by probe on Oct 16 is shown.
- ➔ Five line graphs correspond to movement of probe five times from Guduvanchery to airport during entire day.
- ➔ Vehicle were observed to have greater speed during trip 1 which took place in early morning.
- ➔ As the the time elapsed, vehicles had lower speed as seen in trip 2 which is further supported by higher travel time.
- ➔ Such analysis could be used to identify areas and time of congestion
- ➔ It could be observed that the if the average speed of vehicles is lower, it is very likely that the area has congestion problem and vice versa.
- ➔ Such conclusions could further be supported by looking at corresponding value of flow and density



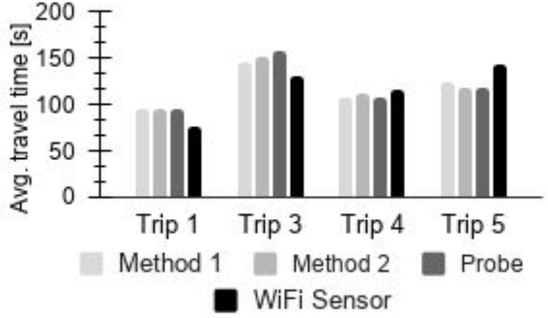


Results

- ➔ Average travel time of probe from Guduvanchery to Chennai airport was compared with travel time estimates made from Method 1 and 2 for the same corridor.
- ➔ It could be seen that the average travel time is maximum during the second trip, which happens during the peak hours during 9:10 and 10:10 AM when many commuters travel along this route.
- ➔ The comparison of the average travel time estimated using proposed methods and corresponding value from independent WiFi detectors on the signalised traffic corridor between Kadeperi and MEPZ signal is done
- ➔ To quantify the precision of estimation using proposed methods Root Mean Square Percentage Error (RMSPE) is used as a precision index.
- ➔ The RMSPE for travel time estimation using Method 1 and 2 is found to be **14.61%** and **16.56%** respectively



Average travel time across corridor



Validation of travel time with WiFi detectors

Conclusion

- ➔ This paper proposed a novel methodology to estimate traffic state on urban arterial using the data recorded by DSRC-based mobile sensors installed in probe vehicles.
- ➔ Modified Edie's generalized definitions have been used to estimate traffic state broadly using two methods, namely- a) Farthest leader and follower and b) Nearest leader and follower configuration.
- ➔ Results suggest that the Nearest leader and follower configuration method better emulate the realtime speed estimation throughout the corridor length.
- ➔ Pearson Correlation coefficient for the speed estimated using Method 1 is found to be **0.8752**. For Method 2, this statistic is found to be **0.9269**.
- ➔ Further study suggests that the method based on the Farthest leader and follower configuration could be used to estimate traffic state variables, namely flow, density with reasonable accuracy.
- ➔ The RMSPE for travel time estimation using Method 1 and 2 on a signalised traffic section is found to be **14.61%** and **16.56%**.
- ➔ Analysis of real-time and historical data would help identify the time and region more prone to congestion with better accuracy.

Discussion

- ➔ Flow and density values could be very sensitive to the small errors in the WiFi data since these estimations are based on the assumption that each moving electronic device is a standalone vehicle. Therefore, there is a need to develop a validation methodology for flow and density.
- ➔ Municipal buses used as probe vehicles often stop at bus stands and traffic signals. Distinguishing these places in the analysis will improve the state estimation results.
- ➔ In addition to this, locations where GPS signal is lost, interpolation of data can be done to calculate distance travelled by the probe.

Research continues in this direction.

Reference

[1] Bhaskar, A., Tsubota, T., & Chung, E. (2014). Urban traffic state estimation: Fusing point and zone based data. Transportation Research Part C: Emerging Technologies, 48, 120-142.

[2] Patra, S. S., Muthurajan, B., Chilukuri, B. R., & Devi, L. (2019, January). Development and evaluation of a low-cost wifi media access control scanner as traffic sensor. In 2019 11th International Conference on Communication Systems & Networks (COMSNETS) (pp. 777-782). IEEE.

[3] Seo, T., Bayen, A. M., Kusakabe, T., & Asakura, Y. (2017). Traffic state estimation on highway: A comprehensive survey. Annual reviews in control, 43, 128-151.

[4] Seo, T., Kusakabe, T., & Asakura, Y. (2015). Estimation of flow and density using probe vehicles with spacing measurement equipment. Transportation Research Part C: Emerging Technologies, 53, 134-150.

[5] Seo, T., & Kusakabe, T. (2015). Probe vehicle-based traffic state estimation method with spacing information and conservation law. Transportation Research Part C: Emerging Technologies, 59, 391-403.

[6] Seo, T., Kusakabe, T., & Asakura, Y. (2015, September). Traffic state estimation with the advanced probe vehicles using data assimilation. In 2015 IEEE 18th International Conference on Intelligent Transportation Systems (pp. 824-830). IEEE.

[7] Montero, L., Pacheco, M., Barceló, J., Homoceanu, S., & Casanovas, J. (2016). Case study on cooperative car data for estimating traffic states in an urban network. Transportation Research Record, 2594(1), 127-137.

[8] Ding, F., Chen, X., He, S., Shou, G., Zhang, Z., & Zhou, Y. (2019). Evaluation of a Wi-Fi signal based system for freeway traffic states monitoring: An exploratory field test. Sensors, 19(2), 409.

[9] Guo, Q., Li, L., & Ban, X. J. (2019). Urban traffic signal control with connected and automated vehicles: A survey. Transportation research part C: emerging technologies, 101, 313-334.