

Urban Traffic State Estimation Techniques Using Probe Vehicles: A Review

Vivek Mehta and Inderveer Chana

Abstract Accurate and economical traffic state estimation is a challenging problem for future smart cities. To curb this problem, fixed roadside sensors are used for traffic data collection traditionally, but their high costs of installation and maintenance has led to the use of probe vehicles or mobile phones containing GPS-based sensors as an alternative cost-effective method for traffic data collection. However, the data collected by the latter method are sparse because the probe vehicles are very randomly distributed over both time and space. This survey paper presents state-of-the-art techniques prevalent in the last few years for traffic state estimation and compares them on the basis of important parameters such as accuracy, running time, and integrity of the data used. The dataset used for the implementation of techniques comes from probe vehicles such as taxis and buses of cities such as San Francisco, Shanghai, and Stockholm with different sampling rates (frequencies) of probes. Finally, it represents the challenges that need to be addressed along with the possible data processing solution.

Keywords Traffic state estimation • Probe vehicles • Data segmentation

1 Introduction

The present demand of traffic state estimation for smart cities requires real-time traffic information which is complete enough and covers a large area of the city. The ultimate aim of getting this information is to build a traffic monitoring system that can be used for the following: 1. Better trip planning. 2. Traffic management. 3. Urban roads and highways engineering, and 4. Urban infrastructure planning.

V. Mehta (✉) · I. Chana
Computer Science and Engineering Department, Thapar University,
Patiala 147004, India
e-mail: vivek_mehta90@yahoo.com

I. Chana
e-mail: inderveer@thapar.edu

Already ongoing approaches for the intelligent transportation systems rely on the real-time data collected by the means of fixed roadside sensors, e.g., loop detectors and video cameras, to detect the traffic state variables such as density, average speed, and travel time [1–3]. However, due to high costs involved in their deployment and maintenance, the authorities cannot cover the entire road network with these devices. Also, as they are fixed they measure only the speed on the spot of their location thus a wider view of the speed is not reflected.

An alternate approach to this static kind of approach is to use GPS receivers which are now coming as embedded in vehicles and mobile phones. This has enabled to get the location and speed dynamically on their path of movement (although it is random). The speed and location updates of the vehicles are known as probes; thus, the corresponding vehicles are known as probe vehicles (taxis, buses, ambulances, etc.). These data probes can be transmitted over a cellular network such as GSM/GPRS to a monitoring center for traffic estimation. As these vehicles cover the entire city, the data are collected over a large area, and due to the low cost of onboard GPS receivers, the overall cost of the system is low. However, this new approach faces some challenges. This paper presents a survey of the techniques which have been used to overcome these challenges. One of the challenges is that the distribution of the vehicles is uneven and incomplete in-between the whole space and time. Figure 1 shows the complete life cycle of data in an intelligent transportation system. As it is shown, firstly the probe measurements are taken with the help of GPS-based mobile phones or probe vehicles. After the collection of raw data, pre-processing techniques such as re-sampling, coordinate

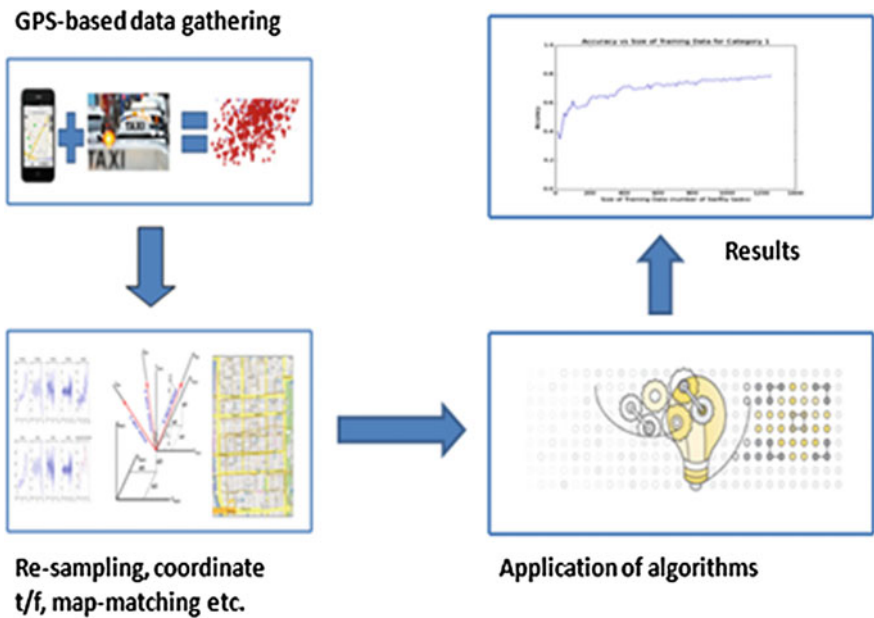


Fig. 1 Data flow in an ITS

transformation, map-matching are applied for suitable application of prediction/estimation algorithms. Coordinate transformation and map-matching are used to represent the collected GPS measurements on a digital map precisely. Re-sampling of data may also be required in case the data are not real and is thus collected in a simulation framework. After this step, one of the different categories of algorithms such as statistical, Bayesian, and machine learning models are applied to get the estimation of traffic state.

In this survey paper, we present a comprehensive study of all the techniques proposed up till now for the estimation of traffic state. This is done by comparing them on the basis of the attributes such as accuracy, running time, sampling frequency of the dataset, and number of vehicles.

2 Survey Approach

This survey has been done with an objective to describe the current state of the art in intelligent transportation research using GPS-based probe vehicles for traffic data collection. This has been done by covering the following aspects.

1. Introduction to GPS-based traffic monitoring system and its advantages over traditional fixed sensor-based technology.
2. Characteristics of a traffic monitoring system.
3. Different approaches that have been used for traffic estimation in the past few years.
4. Challenges that need to be handled in this technology.
5. Existing projects that have implemented this technology.

2.1 Sources of Information and Search Criteria

The research papers used to perform this survey mainly comes from database IEEE explore along with other databases such as Scimedirect (www.sciencedirect.com) and Taylor and Francis (www.tandonline.com).

Search criteria involved using relevant combinations from a set of strings such as GPS-based techniques, traffic state estimation, urban traffic monitoring, and probe vehicles.

3 Brief Review of Existing Approaches

In the past few years, several approaches have been proposed to measure the traffic state variables, for example Bayesian, compressive sensing, aggregation, curve-fitting, statistical, neural-network, and K-NN. They are summarized as follows:

3.1 Bayesian Network Approach

In this approach [4], firstly an expectation-maximization approach is used to learn about the state variables of traffic congestion from a historically available large dataset. Traffic state prediction is done in real time with streaming data. Historical training provides more robustness to the model. In order to make an improvement over another approach of *scaling partial link travel time in proportion with length of the partial links*, it [4] uses density modeling to estimate partial link travel times from link travel times.

3.2 Compressive Sensing Approach

This algorithm is inspired by the idea that applying the method of *principal component analysis* discovers some hidden structures in a large dataset of probe data [5]. Now using compressive sensing, these internal structures are exploited for traffic state estimation. This algorithm has been claimed to outperform KNN and MSSA.

3.3 Curve-Fitting-Based Method and Vehicle-Tracking-Based Method

These two algorithms have been simultaneously proposed by a single author [6]. To support the traffic state estimation of the algorithms, a new method to construct an exact GIS-T digital map has been proposed. On comparing the two algorithms with each other, it has been found that the vehicle-tracking-based method provides higher accuracy but takes more than double the time taken by curve-fitting method.

3.4 Statistical Model

In this model, the total travel time on the road network is considered as the sum of travel time on links and delay at traffic signals and intersections [7]. Trip conditions such as time of the day, season, and weather conditions, and the network characteristics are taken into account by expressing the mean and variance of link travel times and turn delays as functions of explanatory variables in combination with fixed effects for groups of segments.

3.5 Artificial Neural Network

An artificial neural network of 3 layers has been proposed in [8] which takes vehicle's position, link id's time stamps, and speed as the input. In it, the model has been compared with Hellinga's model and has been found to perform better than that probably due to higher number of parameters used than those of Hellinga's model. Mean absolute percentage error was less than 6% and is inversely proportional to traffic demand. However, in [8], real GPS data were not used.

4 Comparison and Analysis of Results of Different Approaches

Table 1 shows the complete summary of results achieved by different techniques applied for traffic state estimation. In total, 7 different techniques are compared on the basis of 7 different parameters. The important parameters are explained as follows:

1. *Accuracy*: It is defined as:

$$Ac = 1 - MAPE$$

where MAPE is mean absolute percentage error and is represented as [6]:

$$\frac{1}{N} \sum_{n=1}^N \frac{|\overline{v_n} - \overline{v_a}|}{\overline{v_a}}$$

Here, N is the total number of estimated values during the experiment, $\overline{v_n}$ is the estimated value of the mean speed and $\overline{v_a}$ is the actual value of the mean speed.

2. *Running Time*: It is the time taken by the algorithm for one cycle of estimation. A cycle is the interval of time for which a new traffic state is estimated each time it gets elapsed (time granularity), for example a new estimation can be made after each interval of 4 min or it may be an interval of 15 min.
3. *Temporal Integrity*: This parameter defines the percentage of the total number of time intervals for which GPS sample points appeared for each link of the network on an average. It is worth mentioning here that the actual temporal integrity can vary from as low as '5%' for some of the links to as high as '90%' for the other links.
4. *Sampling Interval*: It is defined as the time elapsed between two consecutive probe reports. A probe report is the speed and location update sent by the probe vehicle to the monitoring center.
5. *Time Granularity*: The traffic state from the collected data is time when a particular interval of time is elapsed. This time interval is known as the time granularity, for example state can be estimated after each interval of 5 min or after each interval of 15 min.

Table 1 Comparison of techniques used for traffic estimation

Method	Accuracy (%)	Running time (sec)	# links, # nodes	#Probe vehicles	Temporal integrity (%)	Sampling interval	Time granularity (min)
Curve fitting	75.42	34	150000, 17000	16000	68	NA	4
Vehicle tracking	83.08	71	15000, 17000	16000	68	NA	4
Compressive sensing	75	0.827	221, NA	4000	60	30 s-several minutes	15
Naïve KNN	55	<0.1	221, NA	4000	60	30 s-several minutes	15
Correlation-based KNN	40	<0.1	221, NA	4000	60	30 s-several minutes	15
Bayesian	85-90	NA	815, 527	500	NA	1 min	15
ANN	95	NA	NA	Simulated data were used	NA	1 min	NA

In general, the major features of a short-term forecasting system are as follows (originally described by Eleni et al. in [9] and further used by Soufiene et al. in [10]).

1. *Determination of the Scope*: It relates to determining that whether our forecasting model should be implemented as a part of traffic management system (TMS) or a traveler information system (TIS) along with the area of implementation (e.g. freeway, highway, and urban arterials).
2. *Conceptual specification of the output*: Here, we specify the size of the horizon and the step. The forecasting horizon denotes the extent of the time ahead which the traffic state has been predicted for. The forecasting step is the actual time interval upon which traffic state is forecasted and thus gives the frequency of prediction in the forecasting horizon. So, intuitively the larger is the forecasting horizon, lesser will be the accuracy of the model. The shorter is the forecasting step used, the more difficult will be to predict. A 15-min interval of time has been indicated as the best interval by the Highway Capacity Manual (2000).
3. *Methodology used for the modeling of data*: An appropriate selection of the methodology should be made for traffic forecasting. In 1990s, ARIMA models were applied to forecast parameters such as urban traffic volume, bottleneck formulation in a freeway, but ARIMA models have a tendency to concentrate just on means leaving the extremes. Over the last decade, techniques such as artificial neural network, nonparametric regression are being widely used.

The results presented in Table 1 shows that the top-performing techniques are ANN, Bayesian, and vehicle-tracking method having more than 80% of accuracy. In order to get more accurate results, these models should be ensembled into one single model with such credibility factors, so that the resulting ensemble model estimates the traffic state with a greater accuracy. To further improve the accuracy and the execution time of the algorithm, data segmentation approach as mentioned in [11] can be applied.

5 Existing Projects Using Probe Technology

This section discusses the projects that used probe vehicles and mobile phones for getting information about traffic flows.

In California, an experiment named Mobile Century [12] was conducted to demonstrate the feasibility of a traffic monitoring system based on GPS-enabled mobile phones. The experiment used the concept of virtual trip lines [13] as its sampling strategy to collect the measurements and send updates. The experiment was conducted using 100 vehicles running in loop on a 10-mile highway in California and the mobile device used was NOKIA N95. Travel times generated by VTL were compared with those generated by loop detectors, and it was suggested

that ‘VTL measurements are more likely to be closer to the actual velocity observed on the field’ [12].

In Hanoi, Vietnam, Hitachi started a demonstration project [14] in 2011 that collected and processed probe data from 300 vehicles in the first year, 2011, and 800 vehicles in the second year, 2012. The output graphics, data, and other forms of information could be used for the estimation of the traffic state. The accuracy achieved for traffic situation identification was approximately 70%. Another ITS project by Hitachi in the province of Bali, Indonesia, obtained GPS data from 300 vehicles operated by a local taxi company. Travel times, speed for a section of road, and travel time for a choice of route were calculated [15].

In the past decade, the potential of smartphones has been exploited by researchers for carrying out many traffic-related tasks such as road incident detection, traffic crowd-sourcing, and traffic queue length detection [16] gives a comprehensive review of all the endeavors that has been done in this area. After analyzing and comparing the existing systems that exclusively depends on mobile phones, [16] states that it is certainly possible to implement a vehicle monitoring system that provides an adequate performance using smartphone-based sensing especially for a developing country.

6 Conclusion

GPS-equipped probe vehicles have come out to be a very promising medium to collect traffic data as it can cover a larger area of road network as compared to fixed sensors. In this paper, we have exhaustively summarized the latest techniques that exploited this form of data and arrived to the best models among them. Further, we proposed that combining the top-performing models (ANN, Bayesian, vehicle-tracking method) with suitable credibility factors into an ensemble model would result into a more accurate model. In addition, a data segmentation approach can result in an even more accurate model with lesser execution time.

References

1. Bramberger M, Brunner J, Rinner B, Schwabach H (2004) Real-time video analysis on an embedded smart camera for traffic surveillance. In: Real-time and embedded technology and applications symposium, 2004. Proceedings. RTAS 2004. 10th IEEE, IEEE, pp 174–181 (May)
2. Coifman B (2002) Estimating travel times and vehicle trajectories on freeways using dual loop detectors. *Trans Res Part A Policy Pract* 36(4):351–364
3. Kong QJ, Li Z, Chen Y, Liu Y (2009) An approach to urban traffic state estimation by fusing multisource information. *IEEE Trans Intell Trans Syst* 10(3):499–511

4. Hofleitner A, Herring R, Abbeel P, Bayen A (2012) Learning the dynamics of arterial traffic from probe data using a dynamic Bayesian network. *IEEE Trans Intell Trans Syst* 13 (4):1679–1693
5. Zhu Y, Li Z, Zhu H, Li M, Zhang Q (2013) A compressive sensing approach to urban traffic estimation with probe vehicles. *IEEE Trans Mob Comput* 12(11):2289–2302
6. Kong QJ, Zhao Q, Wei C, Liu Y (2013) Efficient traffic state estimation for large-scale urban road networks. *IEEE Trans Intell Trans Syst* 14(1):398–407
7. Jenelius E, Koutsopoulos HN (2013) Travel time estimation for urban road networks using low frequency probe vehicle data. *Trans Res Part B Methodol* 53:64–81
8. Zheng F, Van Zuylen H (2013) Urban link travel time estimation based on sparse probe vehicle data. *Trans Res Part C Emerg Technol* 31:145–157
9. Vlahogianni EI, Golias JC, Karlaftis MG (2004) Short-term traffic forecasting: overview of objectives and methods. *Trans Rev* 24(5):533–557
10. Djahel S, Doolan R, Muntean GM, Murphy J (2015) A communications-oriented perspective on traffic management systems for smart cities: challenges and innovative approaches. *IEEE Commun Surv Tutor* 17(1):125–151
11. Bernas M, Placzek B, Porwik P, Pamuła T (2015) Segmentation of vehicle detector data for improved k-nearest neighbours-based traffic flow prediction. *IET Intell Trans Syst* 9(3): 264–274
12. Herrera JC, Work DB, Herring R, Ban XJ, Jacobson Q, Bayen AM (2010) Evaluation of traffic data obtained via GPS-enabled mobile phones: the Mobile Century field experiment. *Trans Res Part C Emerg Technol* 18(4):568–583
13. Hoh B, Gruteser M, Herring R, Ban J, Work D, Herrera JC, Jacobson Q (2008) Virtual trip lines for distributed privacy-preserving traffic monitoring. In: *Proceedings of the 6th international conference on mobile systems, applications, and services*. ACM, pp 15–28 (June)
14. Okubo T, Yoshioka K, Nakamura A, Taniguchi N (2014) Realizing smart mobility using probe data. *Hitachi Rev* 63(6):359
15. Morioka M, Kuramochi K, Mishina Y, Akiyama T, Taniguchi N (2015) City management platform using big data from people and traffic flows. *Hitachi Rev* 64(1):53
16. Engelbrecht J, Booysen MJ, van Rooyen GJ, Bruwer FJ (2015) Survey of smartphone-based sensing in vehicles for intelligent transportation system applications. *IET Intell Trans Syst* 9 (10):924–935
17. Engelbrecht J, Booysen MJ, van Rooyen GJ, Bruwer FJ (2015) Survey of smartphone-based sensing in vehicles for intelligent transportation system applications. *IET Intell Trans Syst* 9 (10):924–935