

Real-Time Link Travel Time Estimation: Using Buses as Traffic Probes

Rouzbeh Forouzandeh¹, Navid Khademi², and Farshad Hakimpour³

Abstract—Link travel time is one of the most important variables in determining travel routes and start time of journeys. Link travel time is the basis of navigation and routing systems. Time dependent algorithms calculate fastest routes based on link travel time of transportation networks. By increasing use of portable receivers of global positioning systems, researchers are more interested in utilizing the data provided with such equipment in monitoring traffic states. In this research, travel time of arterial links is estimated in real time through the data received from Tehran's transit buses AVL system. Link travel time estimation requires some analyses on spatial and temporal data. In this regard, the Holt-winters analysis is used for a short term travel time prediction. Traffic signal timing data are also involved in the computations. Three test vehicles are utilized to provide data required in validating proposed method and also estimating parameters of the model. Finally, estimated travel times are compared with the results of the baseline method. The RMSE of the proposed method indicates the accuracy of the travel time prediction provided by the proposed method. This accuracy indicates the efficiency of the proposed method in the link travel time estimation, which could be a promising outcome for using buses as traffic probes.

Keywords—Bus probes, Link travel time, AVL system, Holt-Winters time series analysis.

I. INTRODUCTION

AS schedules become more accurate, the importance of planning is more evident than ever. A large part of people's time is wasted in crowded transportation networks that not only leads to increased fuel consumption, but also is a treat to community health. Many location based services are designed to help people with their travel plans. While accessing up to date and reliable information of traffic state throughout huge transportation networks is an integral part of such services, developing novel and automatic methods of traffic monitoring is necessary in such a situation. Some methods have been already used for arterial link travel time estimation. Taylor et al. has divided the current methods into two main categories [¹].

The first category estimate link travel time based on fixed detectors. Such methods are known as site base methods. The second category of are those in which link travel time is estimated through utilizing the position of moving vehicles in traffic flows. These methods are based on the data provided by the probe or floating vehicles. Probe based methods can be categorized as GPS based, Bluetooth based, and RFID based methods.

Probe vehicles data have already been used in estimating link travel time. Van Zuynen et al. modeled delays at traffic signals in various traffic conditions by using probe vehicle data [²]. They estimated delays at traffic signals by calculating parameters of delay distribution functions through rebuilding traffic light queues. In their study, the Genetic Algorithm method along with the maximum likelihood estimator was employed to determine the optimum parameters of the delay distribution functions at traffic signals.

Hofleitner et al. introduced a new approach for arterial link travel time estimation by using GPS data gathered from taxicabs of California [³]. They divided link travel time into two parts including travel time of free flow speed and delays related to traffic signals. The researchers modeled distribution function of travel time and then estimated the parameters by using GPS data received from taxi fleet. They assumed that traffic flow has a uniform distribution while modeling link travel time distribution. The results indicated that Hofleitner's proposed approach has a better performance compared with older methods.

Along with cars, buses have also been used as traffic probes. Nobuhiro et al. used bus probes to investigate the travel time variability [⁴]. They acquired travel time distribution functions by analyzing AVL data of transit buses. There are also some other researches devoted to bus probes [⁶], but they rarely address the link travel time estimation.

This paper proposes a novel method using transit buses for monitoring traffic state of arterials. In many cities, transit buses are equipped with AVL¹ systems by which geographical position of buses is measured and recorded over time. This data can be used for the real time monitoring of traffic states. The proposed method in this paper will be introduced in the following sections. Models and results will also be discussed in the subsequent sections.

Rouzbeh Forouzandeh¹ is MSc student in GIS, Department of Geomatic engineering, University of Tehran, Iran (corresponding author's phone: 00989370862917; e-mail: roozbehforouzandeh@gmail.com).

Navid Khademi² is with the Department of Civil Engineering, University of Tehran, Iran (e-mail: Navid.khademi@ut.ac.ir).

Farshad Hakimpour³ is with the Geomatic Engineering Department, University of Tehran, Iran (e-mail: fhakimpour@ut.ac.ir).

¹ Automatic Vehicle Location

II. PROBLEM DEFINITION AND RESEARCH OBJECTIVE

Transit buses are used as probe vehicles in this research. The reason for choosing transit buses as traffic probes is that in many cities, as it is true in Tehran, transit buses are the only vehicles equipped with AVL systems. In addition, due to the fact that buses are usually in main arterials, they can be good candidates to probe traffic state.

In this research, arterial link travel time is estimated for time intervals using data provided by the AVL system of Tehran's transit buses. Arterial links are parts of arterial networks that are located between two consecutive traffic signals. Links' travel time is the average amount of time that vehicles spend between entering and leaving a link in a specific time interval. Time intervals in which link travel times are estimated, may vary between 5 to 30 minutes and is adopted based on the approximate rates of observations (number of bus positions measured by AVL system over time).

Tehran's transit buses AVL system is programmed to record bus positions every 2 minutes. These positions are measured by onboard positioning devices installed on buses and are sent to a central server in which these positions and their corresponding times are recorded. In addition to this time based positioning, times and positions of buses are measured and recorded when bus doors are opened and closed. This extra data is used while determining the bus stop dwell time.

The ultimate objective of this research is to develop a method for estimating link travel time of cars in real time by using historical and streaming data provided by Tehran's transit buses AVL system. Link travel time is divided into two separate parts: (i) delays related to traffic signals, and (ii) travel time of free movement. The effects of traffic signals and bus stops are elaborately considered in this research.

III. ESTIMATING BUS STOP DWELL TIME

According to what was stated in previous sections, it is straightforward to estimate the period of time in which a bus is completely stopped at a bus stop. Actually, it can be fairly assumed that a bus is fully stopped as long as its doors are opened and begins moving when the doors are closed. While both events (doors opening and closing) are recorded by the AVL system, it is not a big deal to estimate stop time of a bus at a bus stop. But bus stop dwell time is not only limited to stop time of a bus. In other words, deceleration and acceleration of a bus during entering and leaving a bus stop is also a contributing factor in the bus dwell time.

In this research, the function of bus stop dwell time includes both aforementioned delays and also considers traffic state as a factor in bus dwell time. Each bus stop has its own functions for estimating bus dwell time in different traffic states. There is an assumption in modelling bus stop dwell time which is:

Assumption 1: Before and after a bus stop, a bus has a constant speed of v_{out} and has a constant acceleration while entering and leaving the bus stop.

v_{out} may be different for different buses at different time intervals, but acceleration and deceleration is considered constant (Fig. 1).

A bus is called a *hypothetical bus* if it does not stop at bus stops. A hypothetical bus travels at the speed of v_{out} or waits at a traffic signal. v_{out} is called *stop outside speed* in this paper. Stop outside speed in link l at time t is indicated by $v_{out}^{l,t}$. Indices l and t are hereafter removed from our equations for the sake of simplicity.

As it is illustrated in Fig. 1, the bus starts with a constant deceleration a_d at T_1 and stops at T_2 . The bus leaves the stop at T_3 and continues its movement with a constant acceleration a_a until T_4 . As a result, the bus speed can be calculated as follows:

$$v(t) = \begin{cases} v_{out} & \text{if } t \leq T_1 \\ v_{out} - v_{out} / (T_2 - T_1) \cdot (t - T_1) & \text{if } T_1 \leq t \leq T_2 \\ 0 & \text{if } T_2 \leq t \leq T_3 \\ v_{out} / (T_4 - T_3) \cdot (t - T_3) & \text{if } T_3 \leq t \leq T_4 \\ v_{out} & \text{if } T_4 \leq t \end{cases} \quad (1)$$

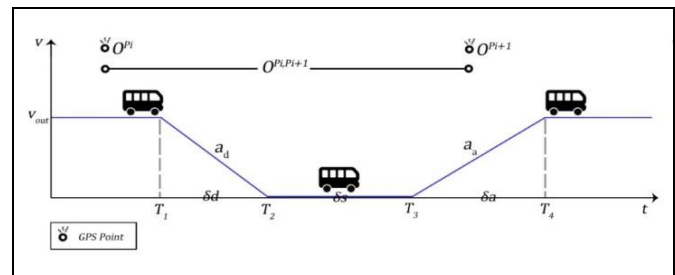


Fig. 1 Bus acceleration and deceleration during entering and leaving a bus stop.

T_1 and T_4 must be determined so that δ_d and δ_a can be specified. If T_1 and T_4 are determined, δ_d and δ_a will be:

$$\delta_d = (T_2 - T_1) / 2 \text{ or } T_1 = T_2 - 2\delta_d \quad (2)$$

$$\delta_a = (T_4 - T_3) / 2 \text{ or } T_4 = 2\delta_a + T_3 \quad (3)$$

Estimated values of δ_d and δ_a for n different traffic states are indicated by $\hat{\delta}_d^{c_k}$ and $\hat{\delta}_a^{c_k}$ ($k \in \{1, 2, \dots, n\}$). Traffic states are categorized by v_{out} . v_{out} , T_1 and T_4 are not known in this step. These variables will be estimated by an iterative method which will be discussed in the following sections.

IV. ESTIMATING v_{out} AND THE QUEUE LENGTH IN PREVIOUS INTERVALS

v_{out} and queue length can be estimated for the previous time intervals provided that AVL historical data and signal

timing data is available. As it will be indicated in the next sections, estimating link travel time in the past time will be the basis for estimating link travel time in real time and in the future.

Link travel time is estimated through solving observation equations which are formed based on historical data. Equations are formed based on the observations (measurements) of AVL system and signal timing information. To form the equations in terms of some elementary variables and parameters, queueing theory in hydrodynamic traffic theory is utilized [3]:

$$n(t) = \begin{cases} n_0 + q \times t & \text{if } t < t_r \\ n_0 + q \times t - s \times (t - t_r) & \text{if } t \geq t_r \end{cases} \quad (4)$$

Where, n_0 is the initial queue length when the signal switches from green to red, q (vehicle / sec) is the link arrival flow, s (vehicle / sec) is the leaving flow during green phase, and t_r is the length of red phase.

In this research q and s are replaced with q_m (meter / sec) and s_m (meter / sec). Using these new variables, it will be easier to form the observation equations which are based on spatial units. Considering the aforementioned idea, (4) is changed to:

$$n(T) = n_0 + q_m \times T - s_m \times T^g(T) \quad (5)$$

Where, T is the time passed from the start of a time interval and $T^g(T)$ is the sum of the green phase between the start of the time interval and T . Equation (5) is the basis of spatial calculations in this research.

AVL data are used in forming observation equations. Two consequence AVL records P_1 and P_2 are considered as an observation O^{P_1, P_2} . Fig. 2 indicates a special case of observations. Two observation equations are formed based on the observation:

$$x(t_1) - x(t_2) = L_w + L_m \Rightarrow x(t_1) - x = s_m \times (T^g(t_2) - T^g(t_j)) + v_{out} \times (T(t_j) - T(t_1)) - \delta^c(O^{P_1, P_2, t_1, t_j}) \quad (6)$$

$$x(t_1) - L_w = n(T(t_j)) \Rightarrow x(t_1) - v_{out} \times (T(t_j) - T(t_1)) - \delta^c(O^{P_1, P_2, t_1, t_j}) = n_0 + q_m \times T(t) - s_m \times T^g(t_j) \quad (7)$$

$T(t)$ is the transformation function that maps regular time to the time frame whose origin is start of the current time interval and $\delta^c(O^{P_1, P_2, t_1, t_j})$ is the bus stop dwell time between t_1 and t_j .

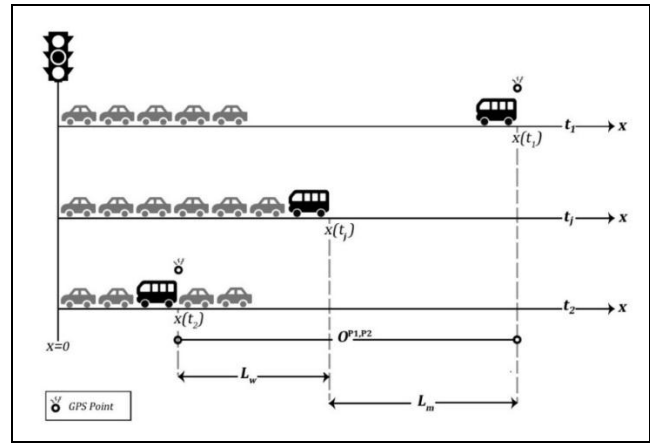


Fig. 2 A special case of AVL observations (both AVL points are in the same link).

There are some assumptions and definitions that must be introduced before forming and solving observation equations. A time interval $\Delta t^{t,d}$ is characterized by the ordered pair (t, d) , where t indicates the t^{th} time interval in day d . Time intervals with similar t are called similar time intervals. $\Delta T_s^t = \{\Delta t^{t,d} \mid d \in D_{\text{historical}}\}$ is the set of all similar time intervals at t^{th} time interval of different days.

Assumption 2: q_m and s_m are constant during a time interval in a link. In addition, q_m and s_m are equal in all similar time intervals.

Assumption 3: v_{out} is constant during a time interval.

4.1. Observation Equations for a Link in a Time Interval

Equations (6) and (7) are only related to a special case of observations. Actually, there are 8 different cases which model different observations and corresponding equations. Since the purpose of forming the equations is to estimate q_m, s_m, v_{out} , and n_0 for all links in all of the previous time intervals, we form and solve the equations for all of the observations. Considering the number of variables and complexity of the equations, solving such equations simultaneously is not an easy task. Our proposed approach is to form the equations, separately, for each link using all the observations received in similar time intervals and based on the approximate values of variables of adjacent links. By this method, observation equations for link l and for similar time intervals ΔT_s^t is formed simultaneously (E.g. All the AVL observations recorded in some special days within 11:30 to 11:45 am).

When n is the number of similar time intervals, the number of unknown variables is $2n + 2$. Least squares estimator is used in this research to solve simultaneous equations. Considering the complexity and discontinuity of the

simultaneous equations, an optimization method is utilized by the Genetic Algorithm method.

4.2. Estimating Initial Values of Variables

As it was stated in the previous sections, it is necessary to have approximate values for adjacent links at the moment of forming the simultaneous equations. For this reason, it is assumed that q_m, s_m, v_{out} and n_0 in adjacent links at t , have the same amounts at $t-1$. Considering relatively light changes in this variable over time, this assumption is fairly acceptable for acquiring some approximate (initial) values.

4.3. Iterative Calculation of Bus Stop Dwell Time and Variables of Bus Travel Time

It can be understood from (6) and (7) that the observation equations are based on the bus stop dwell time and variables, and parameters of travel time (q_m, s_m, v_{out} and n_0). Since the parameters of bus stop dwell time is estimated based on all the observations with a similar traffic state, and moreover, the parameters and variables of travel time are estimated based on the observations received at the similar time intervals, it is not possible to estimate these parameters simultaneously. Therefore, it seems to be inevitable to estimate these parameters in a sequential manner.

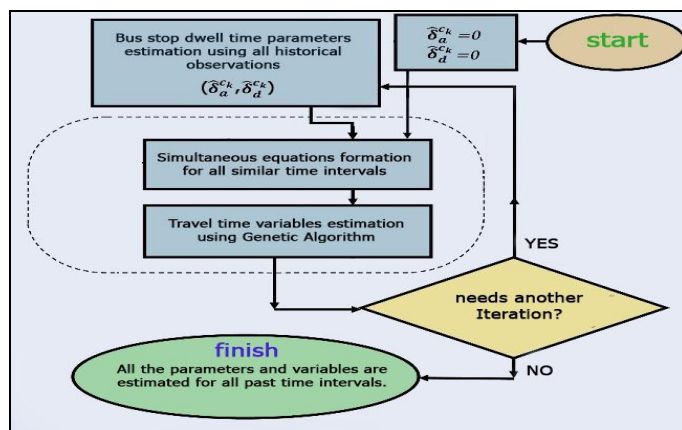


Fig. 3 Process of the iterative calculations of bus stop dwell time and variables of bus travel time.

In addition, the initial values estimated for each group of parameters and variables are only an approximation of their true values. In this situation, it is a good idea to improve the accuracy of the parameters by some iterative calculations. For this purpose, effect of acceleration and deceleration is ignored in dwell time at stops ($\delta_d^{c_k} = 0, \delta_a^{c_k} = 0$ and $k \in \{1, 2, \dots, n\}$). By this method, the initial values of dwell times at stops are determined. Afterward, the variables of link travel time (q_m, s_m, v_{out} and n_0) are estimated. Since some approximate values are taken for the dwell time at stops, the estimation of link travel time parameters has not reached its maximum precision. While some values are now available for parameters of link travel time, parameters of bus stop dwell time can be recalculated more precisely. This iteration

continues while all the parameters reach their maximum precision. Fig. 3 indicates the iterative method.

V. SHORT TERM PREDICTION OF LINK TRAVEL TIME OF BUSES

Link travel time is estimated for the past time by the methods introduced in previous sections. While q_m and s_m are the same in all the similar time intervals, value of these variable are also known in real-time. $v_{out}^{l,t}$ and $n_0^{l,t}$ must be estimated in real-time or near future. One may argue that if simultaneous equations are formed for all the observations received in the current time interval, $v_{out}^{l,t}$ and $n_0^{l,t}$ can be estimated in real-time. However, v_{out} and n_0 for adjacent links ($v_{out}^{l\pm i,t}$ and $n_0^{l\pm i,t}, i > 0$) are also in the equations which make it impossible to estimate $v_{out}^{l,t}$ and $n_0^{l,t}$ by a limited number of observations.

The time series analysis is the proposed solution for this problem. Having the time series of $v_{out}^{l,t}$ and $n_0^{l,t}$ for all previous time intervals, we predict these variables for the future time. Time series must be analyzed to be used as a prediction tool. Different methods exist and are popular for analyzing time series, among them, the Holt-Winters analysis is used in this research [5]. There are some reasons for using the Holt-Winters analysis: (1) it is capable to model seasonal series, (2) recent data is more effective due to its exponential weighting method, (3) it is proved that Holt-Winters analysis has good results in short term predictions, and (4) it is easy to be implemented. It is worth mentioning that in the research, “days” are considered as “seasons” in the time series.

VI. ESTIMATING LINK TRAVEL TIME OF CARS

All the parameters and variables required to estimate link travel time of buses have already been estimated in previous sections, but the ultimate objective of this research, which is the calculation of the link travel time of cars, remained unanswered.

While $v_{out}^{l,t}$ represents bus speed in a link, $v_c^{l,t}$ indicates car speed in link l at time interval t . It has been indicated in previous researches that there is a polynomial relationship between speed of buses and cars [6]. By comparing bicubic, quadratic, and linear cases, a quadratic relationship is finally chosen in this research as followings:

$$v_c^{l,t} = a(v_{out}^{l,t}) + b(v_{out}^{l,t})^2 + c \quad (8)$$

Another assumption should also be considered in this step:

Assumption 4: Delays at a traffic signals are similar for cars and buses.

That is, delays in traffic light queues do not seem to be related to the type of vehicles (buses, cars, etc.). So far, parameters and variables, which are required for estimating link travel time, are calculated; however, the question is: “how

these variables are utilized?” According to Van Zuylen’s study [2], link travel time is calculated by the following equation:

$$TT_{car} = E \{w |t\} + (L - n_0) / v_c \quad (9)$$

Where, L is the length of the link and $E \{w |t\}$ is the mean value of delays at a traffic signal. This variable is calculated based on q_m , s_m and n_0 . Interested readers may refer to Van Zuylen’s article [2] for more technical details.

VII. IMPLEMENTATION AND RESULT

Some implementation details and important results are presented in this section. The data used in this research are obtained from the AVL system of Tehran’s transit buses serving in lines #8803 and #3011. Data was recorded from December 5, 2014 to December 11, 2014. The studied links (3 links) are located in the Motahari Street in Tehran, the capital of Iran. The lengths of the links are 278, 760, and 312 meters and totally two bus stops are located at these links. Link travel time was measured on December 9, 2014 directly by test vehicles in order to evaluate the proposed method. Handheld GPS devices were used in the test vehicles for recording link travel time of cars. The positions of the test cars were recorded at the rate of 1 second.

The length of the time intervals is chosen to be 15 minutes considering the average number of AVL observations received over time. Link travel times are estimated as the historical link travel times, which form the time series. Fig. 4 indicates the time series of v_{out} for the second link in addition to some measures of prediction accuracy resulted from the Holt-Winters analysis.

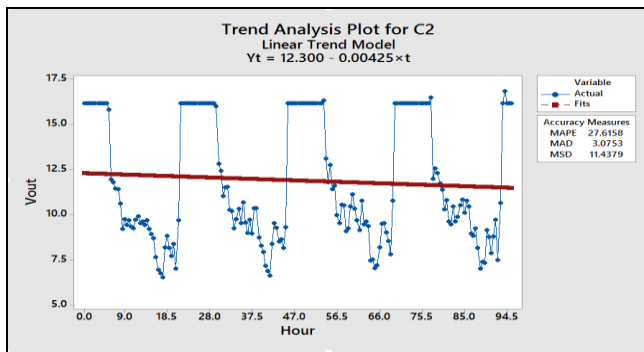


Fig. 4: v_{out} time series for the second link.

The results of the proposed method are compared with the results of the *baseline* method which is known as a basic method for the link travel time estimation [3]. Table 1 shows the RMSE of both methods for some random time intervals. The RMSE of the proposed method provides a 30 percent improvement (on average), compared with the baseline method. The results demonstrate the efficiency of the proposed method in the link travel time estimation.

TABLE I
PRECISION OF REAL-TIME LINK TRAVEL TIME ESTIMATION

Time interval	Base-line method RMSE	Proposed method RMSE
9:00 to 9:15	33.4	15.5
10:45 to 10:30	32.0	18.8
16:00 to 16:15	27.2	22.9
19:30 to 19:15	29.6	20.9
20:00 to 19:45	37.6	16.2

REFERENCES

- [1] M. A. P. Taylor, P. W. Bonsall Young W (2000) Data on travel times. In, Understanding Traffic Systems. Data, Analysis and Presentation. Ashgate Publishing Ltd., 197-206. and W. Young, *Understanding Traffic Systems: Data, Analysis and Presentation*, Ashgate Publishing Ltd, 2000, pp. 197-206.
- [2] H. Van Zuylen, F. Zheng, and Y. Chen, “Using Probe Vehicle Data for Traffic State Estimation in Signalized Urban Networks” in *Traffic Data Collection and its Standardization*, New York: Springer, 2010, ch. 8, pp. 109-127.
- [3] A. Hofleitner, R. Herring, and A. Bayan, “Arterial travel time forecast with streaming data: A hybrid approach of flow modeling and machine learning,” *Transportation Research Part B: Methodological*, 46(9), pp. 1097–1122, 2012.
- [4] U. Nobuhiro, F. Kurauchi, H. Tamura, and Y. Iida, “Using Bus Probe Data for Analysis of Travel Time Variability,” *Journal of Intelligent Transportation Systems*, 13(1), pp. 2-15, 2009.
- [5] M.G. Kendall, and J.K. Ord, *Time-series*, London, Edward Arnold, 1990.
- [6] W. Pu, J. Lin, and L. Long, “Real-time estimation of urban street segment travel time using buses as speed probes,” *Transportation Research Record: Journal of the Transportation Research Board*, 2129, pp. 81-89, 2009.