

Review on Urban Travel Time Estimation Method

Liuren Wang^{1, a}

¹Room 32-2-902, E3 Zone, Quanshui Renjia, Ganjingzi District, Dalian, Liaoning Province, P. R. China

^a20954323@qq.com

Keywords: Travel Time, Prediction Algorithm, Guidance Model

Abstract. Travel time is an important measure of traffic travel service level, and is also an important travel index which can be directly perceived by the travelers. Good calculation application of travel time is an important support for road traffic management, traffic information service and route guidance, and is also an important data support for urban traffic planning. This paper summarizes the relevant parameters and indicators of road travel time, the calculation and distribution principle of the travel time to predict the approximate travel time between any two points of the city, providing a theoretical basis for urban traffic guidance model.

Introduction

Researches on travel time analysis, prediction, reliability and so on are the key issues and links of the realization of the intelligent transportation system construction, such as traffic flow guidance, line guidance system and the decision of the travelers.

According to the research of relevant travel time distribution principle of LPR data, the following four aspects are focused: the relevant parameters and the indicators, the calculation of travel time and the travel time distribution principle.

Related Parameters and Indicators of Travel Time

The Traffic Flow Parameters. *Traffic flow parameters* are physical quantities used to describe and reflect traffic flow. Common parameters include flow, speed, occupancy, etc. Reference [5] gives a more detailed description.

(1) *Traffic volume*: Traffic volume is the number of vehicles that pass a certain point or cross section of a road within a unit time. When the traffic volume exceeds a certain preset threshold value, it is considered to be crowded. However, the same value may correspond to two distinct traffic transport operating states, so the traffic volume is generally not used alone, but needs to be combined with other parameters.

(2) *Speed*: There are many concepts of speed in the road system, including:

a. *Instantaneous speed*, is the instantaneous velocity of a vehicle at a time or through a location, can be used as a basis for traffic control and road design;

b. *The average travel speed* is the value using the length of the road divided by the total time under the premise of containing the parking delay of a vehicle in the process of travelling, when the single vehicle passes a certain length of road. The average travel speed can better reflect the operating state of the vehicle during a specific time period or a specific road. With the relative exclusion of parking delay time is the average running speed index which is less used.

c. *The time average speed* is the arithmetic average of all vehicle speed to pass through a certain location cross section within a time period of observation, reflecting the traffic flow operation state at a particular observation location.

d. *The interval average speed* is the average value of the velocity distribution of all the vehicles passing through a particular length of road at a particular moment, reflecting the traffic flow operating state of a particular road space.

(3) *The time occupancy ratio*: refers to the ratio of time that the detector is occupied by vehicle body and the total detection time during a specific observation period. The time occupancy ratio can better reflect the road traffic operating state. In addition, there is the space occupancy ratio, but because of the direct access to the space occupancy ratio data is very difficult, which is generally less used in the actual project.

The Evaluation Index of Travel Time Prediction Effect. According to reference [1] and [2], the evaluation indexes of predictive effect for travel time include (but not limit to):

Average relative error:

$$MRE = \frac{1}{N} \sum_r \frac{T_r - T_p}{T_r} \text{ Eq.1}$$

Average absolute relative error:

$$MARE = \frac{1}{N} \sum_r \left| \frac{T_r - T_p}{T_r} \right| \text{ Eq.2}$$

Maximum absolute relative error:

$$MAXARE = \max \left\{ \left| \frac{T_r - T_p}{T_r} \right| \right\} \text{ Eq.3}$$

The mean square root of relative error sum of squares:

$$\sqrt{\frac{1}{N} \sum_r \left(\frac{T_r - T_p}{T_r} \right)^2} \text{ Eq.4}$$

Equalization coefficient:

$$EC = 1 - \frac{\sqrt{(T_r - T_p)^2}}{\sqrt{T_r^2} + \sqrt{T_p^2}} \text{ Eq.5}$$

Where T_r represents the actual measurement value; T_p represents the prediction value; N represents the number of forecast samples.

Travel Time Reliability Evaluation Index. Travel time reliability evaluation index can be divided into the following categories:

(1) *Evaluation index based on statistics or bias* refers to using travel time statistics of variance, standard deviation, coefficient of variation to evaluate the reliability/unreliability. The formulas include:

a. Variance:

$$V(t) = \frac{1}{n} \sum_{i=1}^n (t_i - \bar{t})^2 \text{ Eq.6}$$

b. Standard deviation:

$$SD(t) = \sqrt{V(t)} \text{ Eq.7}$$

c. Coefficient of variation:

$$CV(t) = \frac{SD(t)}{\bar{t}} \text{ Eq.8}$$

d. Percent variation:

$$PV(t) = CV(t) \times 100\% \text{ Eq.9}$$

Where t ; represents the travel time spent at the (t) trip within a certain time period on the specified road, n denotes the number of trips during that time period, \bar{t} represents the mean travel time. Compared to the variance or standard deviation, the coefficient of variation and percent variation eliminate the effects of fluctuations in the path-length to the travel time; can more clearly analyze variation characteristics of travel time.

(2) *The Index Based on Buffer Time.* The index based on the buffer time is close to the actual experience of the traveler, which can effectively measure the uncertainty of the highway network, and help travelers to make reasonable travel decisions under the conditions of uncertainty.

a. Buffer index:

Buffer index is developed by the Texas Institute of Transportation, which reflects the more time that the traveler spends than the average travel time in order to arrive at the destination on time with the 95% probability. The formula is:

$$BI(t) = \frac{t_{(95\%)} - \bar{t}}{\bar{t}} \times 100\% \text{ Eq.10}$$

Where, $BI(t)$ represents travel time buffer coefficient, $t_{(95\%)}$ is the 95% of the travel time on the road within a certain period, \bar{t} is the average travel time on the road within a certain period.

b. Florida algorithm:

Florida algorithm is developed by the Florida Department of Transportation, it is widely used in the evaluation of traffic operating state of Florida, which defines the travel time reliability as "the probability when the travel time of the road is less than or equal to the expected travel time plus an acceptable time delays". Where the expected travel time equals to the median travel time within a certain period; the acceptable delay time can be 5%, 10% or 20% of the median travel time, of which the Texas Transportation Institute recommended 10 %.

$$R(t) = \Pr \{t_i \leq t_m + \Delta t\} \text{ Eq.11}$$

Where $R(t)$ is the travel time reliability; t_m represents the median travel time for a certain period; Δt is the acceptable delay time.

(3) *Index based on delayed trip*

The purpose of the evaluation method is to characterize the impact of contingencies on the trip, focusing on the evaluation of delays caused by long travel time, and the statistics of the severity of unexpected delays; generally the bottom 20% of misery index is used to indicate the unreliability of road network in the order from small to large. The formula is:

$$MI = \frac{M_{(t_i > 80\%)} - \bar{M}}{\bar{M}} \times 100\% \text{ Eq. 12}$$

Where MI is the misery index; $M_{(t_i > 80\%)}$ is the average time at the bottom 20% of the travel time in the order from small to large, \bar{M} is the average travel time of all trips.

The greater the misery index is, the greater the impact of uncertainty, it is necessary to avoid the travel of this period. Delayed trip index is different from the index based on statistics; the index based on statistics could eliminate the effect of path-length on highway travel time, while the delayed trip index focuses on the evaluation of delay caused by the longer travel time relative to accidental factors such as weather factors or special events.

Calculation Methods of Travel Time

If the calculation of travel time is divided from the time line, it can be divided into calculation by the historical and real-time data, as well as the calculation by future travel time. According to the content of references, the former is generally referred to as "travel time calculation", "measured value calculation of travel time" [5], "travel time estimation" [6], and the latter for "the prediction of travel time".

Calculation of Measured Value of Travel Time (Travel Time Estimation). The actual measurement of travel time is calculated according to the collected traffic data to calculate the corresponding vehicle travel time, the specific problems and calculation methods needed to consider are related to the data sources [3][4]. Such as the GPS floating car data, because the instantaneous velocity of vehicle is acquired at a moment, which is needed to be combined with other information

such as the longitude and latitude information, using methods such as velocity integral method to calculate the average travel time of the specific road. For fixed detectors such as loop coils, the estimation of travel time is based on the location traffic parameters. For different data source, people present numerous calculation models, such as in the calculation method of travel time for the location traffic parameters, there are the Federal Highway Administration BPR (bureau of public roads) model and kinds of improved models, as well as flow conservation model based on traffic data with the largest effects and most widely used; the direct substitution method and correlation analysis method based on location velocity data; g factor method, regression analysis method and so on based on the occupancy ratio. This chapter gives a detailed introduction to travel time estimation method under the circumstances of the data source for the location traffic parameters, GPS data and the mix of both.

Travel Time Prediction. By prediction time the travel time prediction can be divided into short-term prediction and long-term prediction. Currently researches on Travel Time Prediction abroad focused on short-term prediction, that is the forecast period is within 15min (also literature said that is within 30min). According to the content the travel time prediction can be divided into route prediction and path prediction. Where the path is connected by many routes, the path travel time is not the total of all the route travel time. Most prediction methods are for route travel time prediction.

There are many kinds of travel time prediction methods. In their classification, different categories are given slightly different by different literatures, for example Huangyang You [5] divided the main methods into four categories of parameter based model, the non parameter based model, artificial intelligence based model and the combination based model:

(1) The principle of prediction algorithm based on parameter model is by constructing a reliable prediction model to predict the future traffic flow parameters during a time period, including the linear models of time series method, index smoothing method and Kalman filtering method, and the nonlinear models of chaos theory and wavelet analysis.

(2) The algorithm principle Based on non parameter model finds out the evolution characteristics of the traffic flow through the statistical analysis of a number of measured data, and based on which forecasting the short-term traffic parameters. Accurate model is not needed for this algorithm.

(3) The principle of prediction method based on artificial intelligence is to realize the prediction of traffic flow with the aid of artificial intelligence technology. The common methods include neural network, support vector machine, etc.

(4) Prediction method based on combination model is to fully play advantages of various prediction models, combining two or more different prediction models for short term traffic parameter prediction, so as to improve the prediction accuracy and enhance the applicability.

Distribution Principle of Travel Time

The main purpose of the research on the distribution principle of travel time is to analyze and establish a reasonable and reliable evaluation model of travel time. Travel time reliability can describe the stability, which means the probability that the vehicle can reach the destination D from the starting point O [5] within the specified time.

Research on travel time distribution is generally divided into urban roads and highways from the road type, and from vehicle type is distinguished into car [4], bus [6] or full-size vehicle [3]. Traditional route travel time reliability evaluation generally assumes that the travel time is subject to the normal distribution. Reference [5] pointed out that in recent years, scholars at home and abroad used reliability analysis models such as normal distribution, lognormal distribution, gamma distributions and Weibull distribution to carry out curve fitting analysis for road travel time probability distribution, and the general conclusion is fitting effect of lognormal distribution is relatively better. Specifically, the expressway, auxiliary road, trunk road, secondary road by fitting effect from good to bad are followed by lognormal, gamma, normal and Weibull distribution; and for the highway, the fitting effect from good to bad are followed by Weibull, log normal, normal and

gamma distribution [5] [5]. Different scholars have different research conclusions about the most suitable travel time distribution principle.

The Normal Distribution. The normal (or Gauss) distribution is the most widely used model to describe the properties of random variables. The traditional route travel time reliability evaluation usually assumes that the travel time is subject to the normal distribution, that is $t \sim N(\mu, \sigma^2)$. The probability density function of the normal distribution is:

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] \text{ Eq.13}$$

Where t represents the route travel time, $t > 0$; μ represents location parameter as the expectation of route travel time; σ is scale parameter, representing the standard deviation of route travel time.

The Lognormal Distribution. In the system reliability analysis, the log normal distribution is more widely used than the normal distribution, because it has a series of shapes, which can better fit the reliability data. The probability density function of the lognormal distribution is:

$$f(t) = \frac{1}{\sigma t \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln t - \mu}{\sigma}\right)^2\right] \text{ Eq.14}$$

Where t represents the route travel time, $t > 0$; μ and σ represent expectation and standard deviation of $\ln t$ (logarithm of link travel time), which is the scale parameter representing standard deviation of travel time. When $\mu \gg \sigma$, the lognormal distribution is close to the normal distribution.

The expectation and standard deviation of travel time can be calculated by the following formulas:

$$E(t) = \exp\left(\mu + \frac{\sigma^2}{2}\right) \text{ Eq.15}$$

$$SD(t) = [\exp(2\mu + 2\sigma^2) - \exp(2\mu + 2\sigma^2)]^{1/2} \text{ Eq.16}$$

Where $E(t)$ represents the expectation and $SD(t)$ represents standard deviation of travel time.

The Basic Steps of Travel Time Probability Distribution Fitting and Testing. The basic steps of probability distribution fitting and testing are as follows:

Step 1 Calculate road travel time of unit distance;

Step 2 Select the unit distance route travel time within a certain time period as a data sample to make statistical histogram and preliminary analysis for probability distribution of travel time.

Step 3 Use various distributions (normal, lognormal, etc.) to be tested for curve fitting of travel time data;

Step 4 Use error sum of squares (SSE) and R-squared (R-square) and other indicators to measure the effect of curve fitting;

Step 5 Use Kolmogorov-Smirnov test (K-S test), χ^2 test (chi-square test) and other hypothesis testing methods to verify the rationality of the distribution.

Acknowledgements

This manuscript is the outcome as the project "Research on Urban Traffic Guidance Model Based on Big Data Mining Technology" (project number: L2015249, general project of the Education Department of Liaoning Province scientific research in 2015, project funds: 60 thousand Yuan).

References

- [1] Jinjin Zhang, Research on Urban Traffic Travel Time Prediction Algorithm and Software [D] (in Chinese), North China University of Technology, 2014.
- [2] Gengxian Zhu, Research on Travel Time Prediction of Neural Network Based on Radial Basis Function (RBF) [D] (in Chinese), Chongqing University, 2004.
- [3] Jiwei Li, Research on City Primary and Secondary Trunk Road Travel Time Estimation and the Prediction Method [D], Jilin University, 2012.

- [4] Yan Zhu, Yanrong Cao and Dosheng Du, Statistical Analysis and Prediction of Urban Expressway Travel Time [J] (in Chinese), Transportation Engineering and Information Journal, 2009 (01): 93-97.
- [5] Kun Chen, Reliability Evaluation and Algorithm of Urban Road Network Travel Time Model Based on Mobile Source Data [D] (in Chinese), Beijing Jiaotong University, 2008.
- [6] Hao Chen, Research on Bus Station Capability Optimization Based on Vehicle Travel Time Distribution [D] (in Chinese), Central South University, 2013.