WILEY | Hindawi

Research Article

Estimation of Urban Link Travel Time Distribution Using Markov Chains and Bayesian Approaches

Wenwen Qin (iD) **and Meiping Yun** (iD)

Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, 4800 Cao'an Highway, Shanghai 201804, China

Correspondence should be addressed to Meiping Yun; yunmp@tongji.edu.cn

Despite the wide application of Floating Car Data (FCD) in urban link travel time and congestion estimation, the sparsity of observations from a low penetration rate of GPS-equipped floating cars make it difficult to estimate travel time distribution (TTD), especially when the travel times may have multimodal distributions that are associated with the underlying traffic states. In this case, the study develops a Bayesian approach based on particle filter framework for link TTD estimation using real-time and historical travel time observations from FCD. First, link travel times are classified by different traffic states according to the levels of vehicle delays. Then, a state-transition function is represented as a Transition Probability Matrix of the Markov chain between upstream and current links with historical observations. Using the state-transition function, an importance distribution is constructed as the summation of historical link TTDs conditional on states weighted by the current link state probabilities. Further, a sampling strategy is developed to address the sparsity problem of observations by selecting the particles with larger weights in terms of the importance distribution and a Gaussian likelihood function. Finally, the current link TTD can be reconstructed by a generic Markov Chain Monte Carlo algorithm incorporating high weighted particles. The proposed approach is evaluated using real-world FCD. The results indicate that the proposed approach provides good accurate estimations, which are very close to the empirical distributions. In addition, the approach with different percentage of floating cars is tested. The results are encouraging, even when multimodal distributions and very few or no observations exist.

1. Introduction

Information of travel time distribution (TTD) is essential for route guidance and advanced traveler information systems. It reveals underlying traffic conditions for measuring travel time reliability, evaluating the level of service of intersection, and detecting the congestion and incident, which can be introduced to improve service reliability performance [1].

Up to now, the increasingly available vehicles equipped with GPS, named as floating cars, provide a significant amount of vehicle trajectory data in urban road networks. Many works have proposed analytical or data-driven models to infer mean link travel time from this data, but yet research on the estimation of link TTD is still evolving, especially when the travel times appear to be multistate features. In reality, the travel times of vehicles traversing an urban link are heavily affected by traffic lights, interaction among vehicles, and conflicting traffic from cross streets [1–5]. The characteristics of interrupted traffic flow would likely result in different clusters of travel times, which reflect how vehicles experienced traffic states on a link. Therefore, rather than to provide the mean value that is generally used in planning applications, a probability distribution, that fully characterizes different components of travel times for retaining maximum traffic state information, is preferred.

In order to estimate the link TTD, a relatively large number of floating cars is needed to generate useful travel times among various errors such as GPS errors, map-matching errors, and estimation errors. Unfortunately, the spatial distribution of floating cars is not uniform. Taxicabs, which have been suggested as the main types of floating cars [6], are used widely to collect travel time information for urban link travel time and congestion estimation [4, 5, 7]. However, taxicabs cover mainly densely populated and

commercially well-developed areas, since these areas may have higher passenger flow and taxicabs are more likely to pick up passengers. For those areas with limited coverage by floating cars, it would raise inquiries about how to utilize the limited data sample from a small number of floating cars to produce an accurate link TTD, even for the multistate feature of travel times.

In this situation, it may be possible to combine real-time traffic information incorporating the current traffic states with prior knowledge from historical link TTDs that provide similar data sequences to the actual distributions to reconstruct the TTD. The objective of this study focuses on developing a methodology to estimate such probability distributions under situation when floating car density is quite low and link travel times may follow multimodal distributions.

The main contributions of this paper are as follows. (1) This research can be regarded as an initial attempt to estimate multistate link TTDs in the context of the sparse and incomplete observations from real-time FCD. The corresponding solutions are given by a Bayesian approach based on particle filter framework, in which the state-transition trends are modeled by historical trends to construct an initial solution providing past travel time patterns, as well as the sampling strategy that focuses on a fundamental issue of the irregular, unsmooth TTDs directly constructed by the only real-time FCD and allows much less real-time information. (2) The results of case study show that multistate link TTD models can be robustly realized by using our proposed approach for different percentage of floating cars, even for the condition with no data available. Moreover, the applications help provide in-depth understanding on the heterogeneity of link travel times that show multistate distributions and imply significantly different state combinations changing with time of day, in which the state converging most of travel times may play a dominant role on the whole link traffic conditions.

2. Literature Review

In review of literature, models on the estimation of travel time distribution (TTD) can be generally divided into two categories, namely, unimodal and multimodal/multistate distributions. Unimodal distributions, such as Weibull, exponential, Lognormal, or normal distributions, are often used for modeling travel time. For example, Hunter et al. [8] proposed a probabilistic model for inferring vehicle path and travel time based on low frequency GPS data in the arterial network. An expectation maximization (EM) algorithm for the proposed model is employed to simultaneously learn the likely paths and the parameters of lognormal link travel time distributions using the path travel time observations.

A recent work of the same group by Herring et al. [9] proposed a statistical approach to capture the evolution of traffic states with a coupled hidden Markov model, since the actual state of current link is unobserved (hidden) and depends on adjacent links (coupled). The authors assumed that travel times between successive links are independent, and only one traffic state on each link can exist in the same period. The authors went on to assume that the link travel time distribution is conditional on its hidden state. Likewise, an EM algorithm is used to iteratively update the parameters of the distributions and the state-transition matrix. The evaluation is done using sparse probe data from taxicabs in San Francisco, CA.

A similar approach was proposed by Westgate et al. [10]. In this work, a Markov Chain Monte Carlo (MCMC) algorithm was used to simultaneously estimate the likely paths, travel times, and parameters of link TTDs from GPS-equipped ambulances. The findings from these studies in general concluded that the lognormal distribution provides a superior fit over the alternative unimodal distributions. The main limitation of unimodal distributions is the assumption that travel time is the result of a single traffic state on a link. Hence unimodal distributions usually provide some poor fit results under complex traffic conditions, such as interrupted flow.

The multimodal travel time distribution comprising two or more distributions outperforms the single distribution by reflecting the coexistence of heterogeneous traffic states simultaneously [11, 12]. One of the most attractive features of the multimodal distribution provides a straightforward relationship for the probability of encountering state (e.g., congestion) and the component distribution accordingly under such state. In [11], a multistate model was proposed by using a mixture of normal distributions to fit travel times on freeways. The authors revealed that different traffic states can exist in the same analysis period such as free-flow and congested states. Besides each component distribution is associated with an underlying state. In the research by Park et al. [12], the authors attempted to quantify the impact of traffic incidents on travel time reliability using multistate (3 states) model. The state of a section on freeways describes the level of congestion. They concluded that the congested state corresponding to the third component distribution becomes more dominant since incidents increase the travel time variability.

While very few multimodal link TTD models using GPS data were proposed up to now, for the estimation of the parameters of multimodal distribution, Hofleitner et al. [13] developed a dynamic Bayesian network for the transition between link states using GPS probe data. A simulation-based EM algorithm is proposed to estimate the transition probabilities and the travel time distribution parameters. Feng et al. [7] integrated signal timing and arterial geometry information with real-time GPS data to dynamically update the parameters of the bimodal urban link TTDs based on MCMC simulation. Ji et al. [1] proposed a hierarchical Bayesian bimodal TTD to capture the interrupted nature of urban traffic flow as well as analyze traffic operations from probe vehicles. A further development using a hierarchical MCMC algorithm for producing the posterior distribution of Gaussian mixture models and the lognormal prior was presented in Li et al. [14]. The authors assumed that real-time GPS samples are distributed in each component distribution. These studies all adapted iterative procedures to estimate the parameters of link TTDs using maximum likelihood estimation and Bayesian approach. However, one limitation of their work is that travel times of probe vehicles are

not readily available for each unobserved state since probe vehicles penetration rate is fairly low. Another limitation of their work is that they assumed the link travel times are independent. In reality, adjacent link travel times may have strong correlations. As an alternative, the multistate nature of travel times and their correlations between adjacent links can be modeled by a Markov chain methodology. Ramezani and Geroliminis [15] proposed a Markov model to estimate arterial route TTD. The model used a Transition Probability Matrix (TPM) constructed from simulated probe data to capture the spatial correlations. However, the approach conducted by Ramezani and Geroliminis [15] needs ample real-time observations conditional on link states to estimate the TPMs, which limits the application under urban areas with missing data (no or few data available).

In this paper, TPMs of the Markov chain between upstream and current links are calibrated from historical travel time observations to alleviate the dependence on the limited data sample. On this basis, the study develops a Bayesian approach based on particle filter framework for estimation of multistate link TTD by combining past information of link TTD represented as an importance distribution with real-time travel time information represented by a likelihood function. The importance distribution, used as an initial solution for approximating the actual distribution, is constructed by the summation of past TTDs conditional on link states weighted by the estimated state probabilities. The likelihood function is chosen as a normal distribution that describes how likely a particle (i.e., sample) tends to follow the real-time observations. As a result, the particles with high similarity are retained and replicated to address the sparsity problem of observations and then can be integrated into a Bayesian algorithm for producing a multistate link TTD.

3. Data Collection and Characteristics

3.1. Data Introduction. Nanjing is the capital of Jiangsu province, lying in the east of China, with over 8,236,000 inhabitants. The city has been rather well equipped with ITS infrastructures for nearly 10 years. Radio Frequency Identification Data (RFID) technology is applied in some main arterials. Currently, more than 0.7 million vehicles, including taxicabs and private vehicles, have been equipped with RFID tags. There are over 7700 GPS-equipped taxicabs to collect vehicle trajectory data as Floating Car Data (FCD). In this paper, FCD are used to evaluate the performance of the proposed approach, as well as RFID being chosen to produce the ground truth distributions compared to the corresponding estimated TTDs from FCD.

FCD vehicles generate a data record every 30 seconds including the vehicle's position, timestamp, speed, and status (with or without passengers). The process of matching the reported locations to road network links from raw data records has been accomplished by a widely used map-matching approach [16]. Since this is not the focus of this paper, we will not cover the detail here. Besides, the impact of different data polling frequency on the required sample size determination is not considered in this paper. In order to eliminate the effects from vacant taxi driving intentionally

slowly, and searching for potential passengers, only FCD from hired taxis are used. As to RFID, several arterials are installed with both RFID sensors and video cameras at upstream and downstream locations of signalized intersections. RFID sensors scan the vehicle plate with high frequency of 20 milliseconds and mainly record plate number, timestamp, and station number when tagged vehicles pass by the RFID stations. More than 80% of vehicles on arterials are detected by RFID sensors [3–5]. Hence RFID can almost obtain the information of all the passing vehicles which constitutes a full-sample data.

As shown in Figure 1, the case road link divided by two consecutive RFID stations (i.e., the stations between 6004 and 6006) is around 1054-meter long. The road link has eight-lanes double way from Caochangmen Street and Qingliangmen Street all along and intersects with many other streets in between. The design speed of the selected studied link is 60 km/h.

For the source of FCD, the 1-month (November 2011) data for weekdays (Monday to Friday) between 7 a.m. and 8 p.m. were taken as the historical database. For both sources (FCD and RFID), data from 7 a.m. to 8 p.m. on Dec. 1, 2011 (Thursday), were used for the comparison and validation. Considering the word limit, we choose the south bound traffic flow to analyze the travel time.

3.2. Characterization of Travel Times. In order to estimate link travel time for individual floating car traversing a link, a travel time allocation method proposed by Sanaullah et al. [17] is applied here because of the simple calculation and high estimation accuracy. Then, mean link travel time can be calculated with an average of travel times for individual floating cars in an analysis interval. Note that an observation from FCD (RFID) is treated as the travel time in each link experienced by a vehicle. Besides, the time interval for estimation is set as 15 min to avoid unavailable observations.

To better understand the differences between the mean value and the probability distribution for explaining the trends or patterns of traffic states under interrupted flow, Figure 2 shows the temporal variation of travel times and the number of vehicles counted from both FCD and RFID in 15-min intervals. Apparently, the development of the mean travel times from FCD and RFID is quite consistent. The mean absolute percentage error (MAPE) is chosen for evaluating the accuracy of the estimated travel time. It turns out that the value of average MAPE is 8.46%, which appears to be quite acceptable.

However, Figure 2(a) shows an intuitive understanding on travel times, which can be clustered into two or more components from the scattered observations in a certain interval. This finding implies that the mean is not a good tracking index for capturing the multistate features of link travel times, once compared to TTD. On the other hand, Figure 2(b) illustrates that the proportions of floating cars are relatively low during time of day and that observations from FCD largely distribute at the high-speed regime (nearly between 25 and 50 km/h). These may likely result in inadequate sample to monitor the remaining travel time states. To
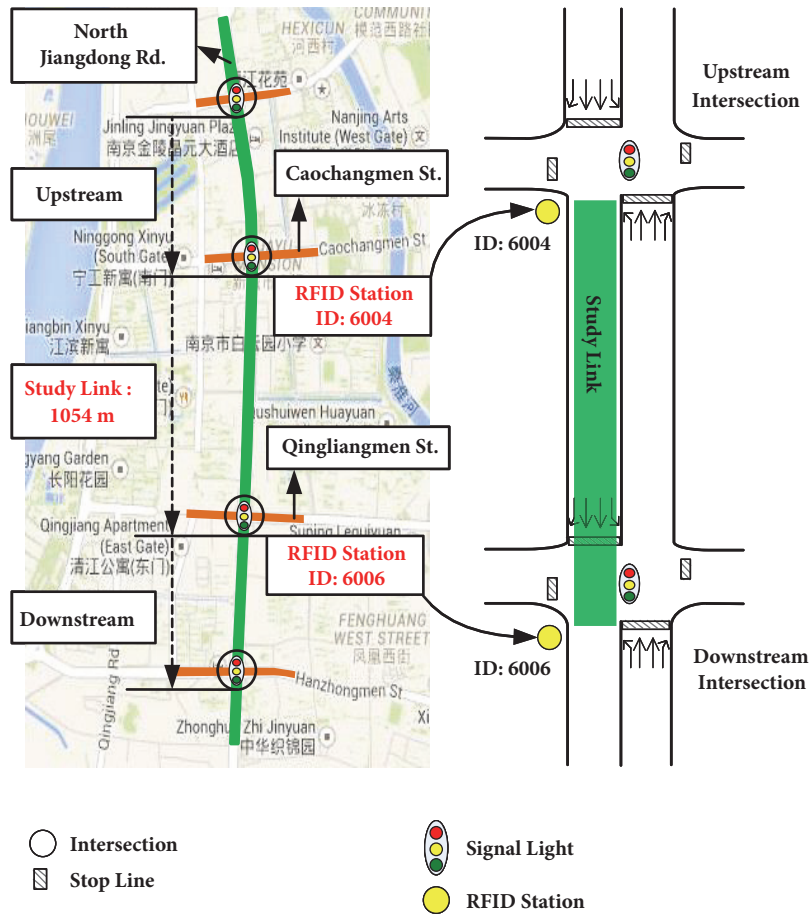
FIGURE 1: Layout of the study link selected in Nanjing, Jiangsu province, China.
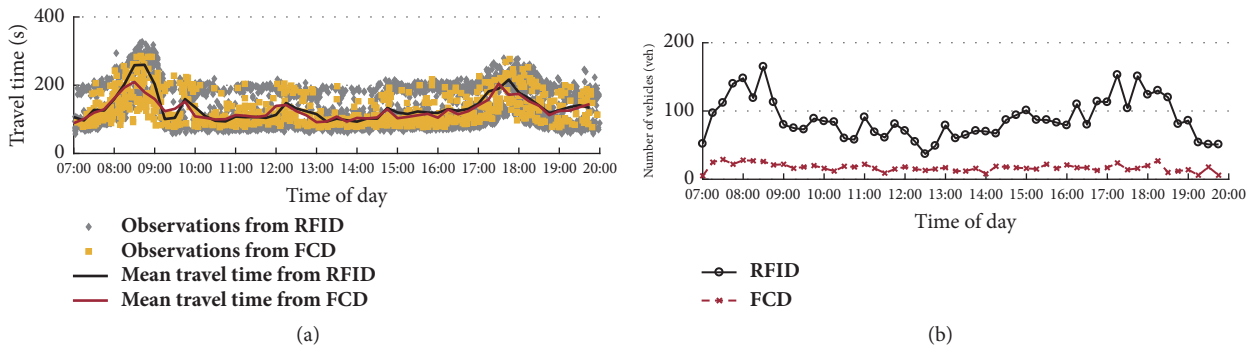

(a)

(b)

FIGURE 2: Variability by a 15-min interval: (a) mean travel times and (b) number of vehicles.

better model multistate TTD, a Bayesian approach based on particle filter framework is then introduced with limited FCD sample.

4. Methodology

The proposed solution framework, consisting of five major parts, is outlined in Figure 3. The state identification is presented first in this section, followed by a description of how the states transfer between two adjacent links. After that, the importance distribution is introduced to produce historical travel time patterns as prior knowledge. The following subsection presents the sampling strategy to generate a certain number of high weighted particles. Finally, the Gibbs sampling algorithm is applied to approximate the posterior distribution of Gaussian mixture model (GMM).

4.1. State Identification. The multistate nature of travel times on urban links is principally attributed to the road geometric features (e.g., roadway alignment, link length, lane width, and number of lanes), interaction among vehicles (e.g., weaving, diverging, and merging), and signal controls. With the results
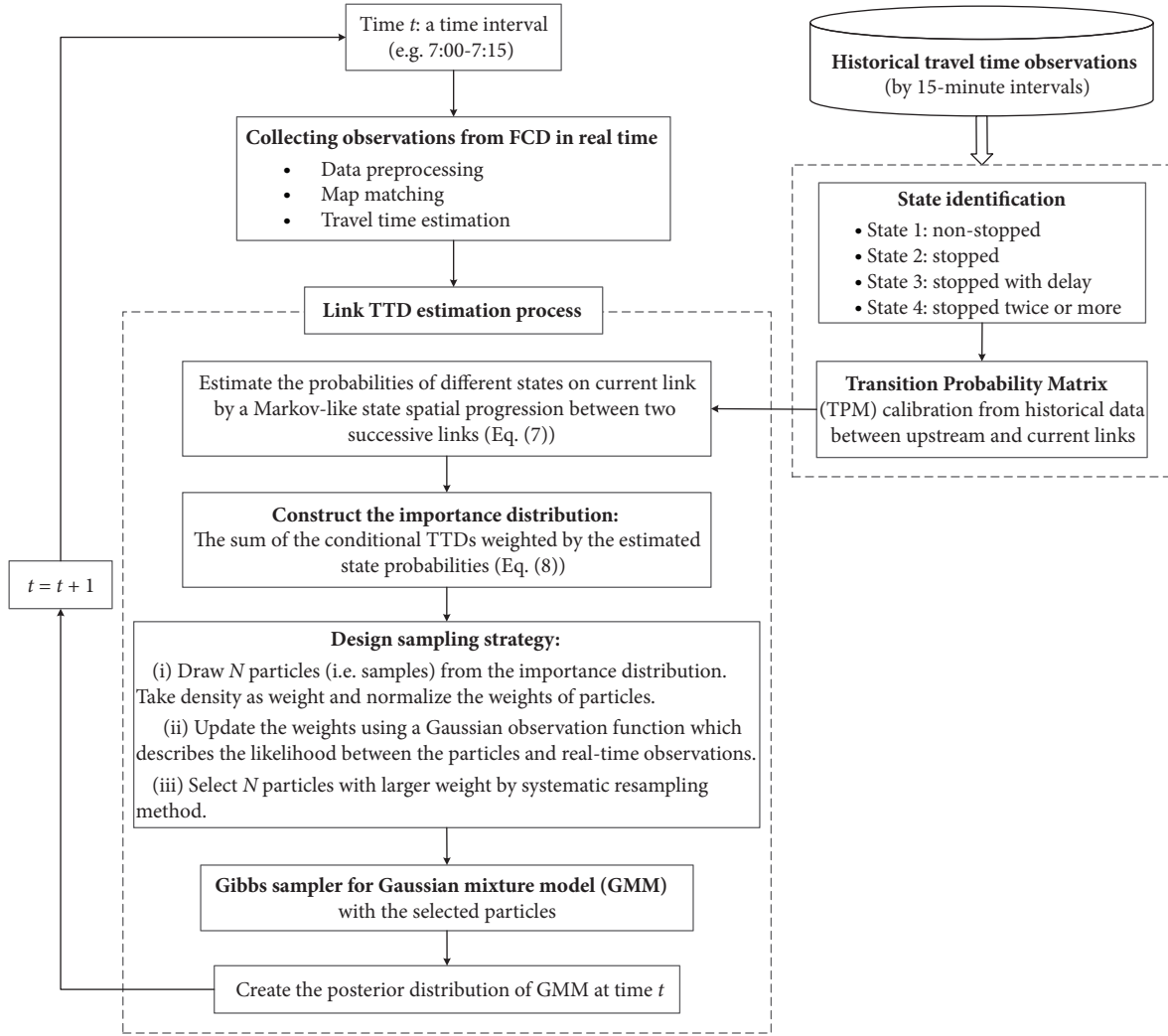
FIGURE 3: Link travel time distribution estimation framework. This figure is reproduced from a shorter version of this work presented at the 97th Annual Meeting of the Transportation Research Board [5].

of the interaction among these factors, four typical traffic states that exhibit travel times on a link experienced by vehicles are shown in Figure 4, and can be defined as follows:

(i) State 1 (nonstopped): vehicles traverse the whole link without stopping.

(ii) State 2 (stopped): vehicles stop at the end of the link for a red light.

(iii) State 3 (stopped with delay): vehicles traverse the whole link with delays caused by the interference among vehicles and signal control.

(iv) State 4 (stopped twice or more): vehicles have to stop twice or more due to the overflow queue at the signalized intersection.

In order to distinguish the heterogeneous traffic states, intuitive traffic engineering rules are applied to quantify state boundaries so each state represents the corresponding cluster of travel times. Here the boundaries between two neighbor states can be approximated as follows, respectively.

(i) The first boundary between state 1 and state 2 is to judge whether vehicles pass the signalized link without stopping.

Considering a free-flow traffic condition, the travel times of vehicles traversing a signalized link directly depend on the signal controls and their distribution appears to be bimodal [1–5, 7], which features two clusters of travel times corresponding to the states of nonstopped and stopped vehicles, respectively. To achieve the state partitions, a two-component mixed normal distribution, rather than the skewed mixture models (e.g., the components are both lognormal distributions), is adopted to obtain the travel times in state 1 and state 2 individually as a tradeoff between computation complexity and explanatory power [18]. Besides, the travel times between 02:00 a.m. and 04:00 a.m. are extracted from FCD for the free-flow condition [19]. Hence, the first travel time boundary (denoted as TT_f), determined by the point of intersection of two normal distributions [20], is stated mathematically as
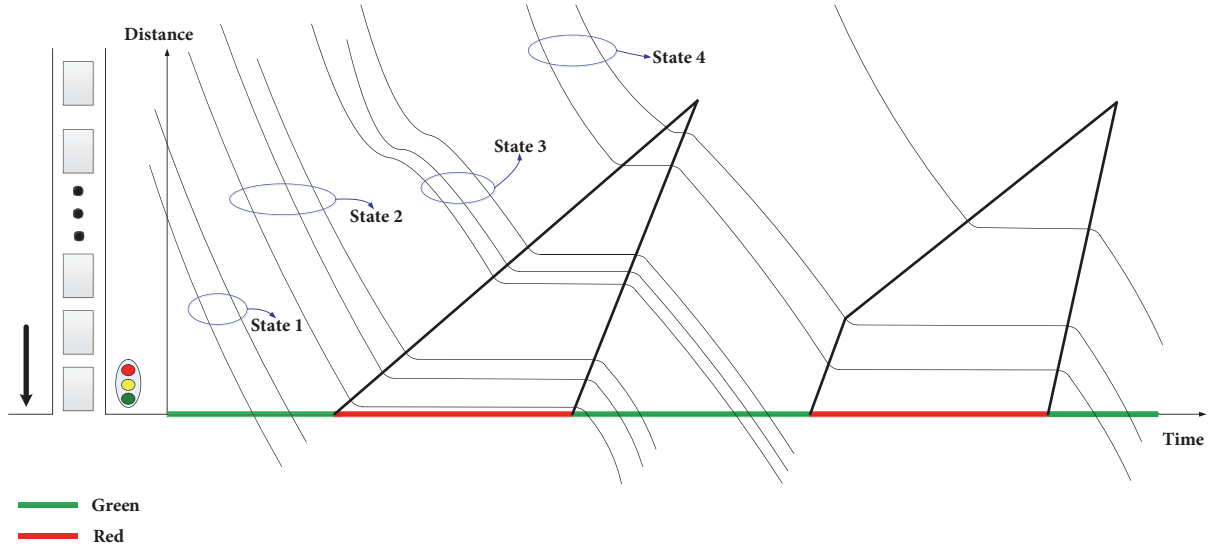
FIGURE 4: Four typical travel time states on urban signalized links.

$$TT_f = \mathrm{N}\left(TT \mid \mu_{f,1}, \sigma_{f,1}\right) = \mathrm{N}\left(TT \mid \mu_{f,2}, \sigma_{f,2}\right), \quad (1)$$

$$\mu_{f,1} < TT_f < \mu_{f,2}$$

where $\mathrm{N}(*)$ is the normal distribution, TT is travel time variable, $\mu_{f,1}$ and $\mu_{f,2}$ are the mean of the first and second normal distribution, and $\sigma_{f,1}$ and $\sigma_{f,2}$ are the standard deviation of the first and second normal distribution.

(ii) The second boundary between state 2 and state 3 is to distinguish whether vehicles only experience intersection signal delay.

Signal delay is the primary factor that influences travel time patterns on an urban link [3]. Unfortunately, in most cases, the signal control plans at intersections are generally unknown. As an alternative, we use the red time for through movement (denoted as TT_{red}) as additional travel times to judge whether vehicles stop at the end of the link. This part of travel times is assumed to be only generated by the signal control, since it makes a major contribution to travel delay rather than traffic flow. Note that the means of $\mu_{f,1}$ and $\mu_{f,2}$ for state 1 and state 2 are closely related to the signal settings, and their difference between nonstopped and stopped vehicles can be used to approximately estimate the red time. Consequently, the second boundary by adding the red time to free-flow travel time (denoted as TT_{stop}) is represented as

$$TT_{stop} = TT_f + TT_{red} = TT_f + \mu_{f,2} - \mu_{f,1} \quad (2)$$

(iii) The last boundary is to judge whether vehicles experience oversaturated conditions.

The oversaturated conditions on urban principal arterials, characterized as extremely low speeds and extensive queuing, is defended as level of service F [21]. In this case, a residual queue is the most likely occurring at the signalized intersection in one cycle time. Hence, the last boundary for oversaturated conditions (denoted as TT_{OC}) where vehicles encounter two or more queues can be initially estimated by link length and 30% of speed limit:

$$TT_{oc} = \frac{L}{0.3 \times V_{limit}} \quad (3)$$

where L is the link length and V_{limit} is the speed limit of the arterial link.

4.2. State-Transition Function Modeling. This section adopts work of Ramezani and Geroliminis [15] to represent the spatial state transition between two adjacent links based on Markov chain framework. The spatial state transition exhibits how vehicles experience multiple traffic states on route's successive links whereas it differs from the traffic congestion propagation. For example, a vehicle that passed the upstream link without stopping would have a certain probability of stopping on the downstream link, and it would have another probability of not stopping on the downstream link. This simple case illustrates that the travel times on the current link are dependent on the upstream link state.

In this paper, there are four traffic states constituting a four by four TPM. The TPM $P_{n-1,n}(t)$ between any pair of successive links n-1 and n at time t is shown as follows:

$$P_{n-1,n}(t) = \begin{bmatrix} p_{11}(t) & \cdots & p_{14}(t) \\ \vdots & \ddots & \vdots \\ p_{41}(t) & \cdots & p_{44}(t) \end{bmatrix} \quad (4)$$

$$\sum_{j=1}^{4} p_{ij}(t) = 1 \quad \text{for } i = 1, \ldots, 4, \ 0 \le p_{ij}(t) \le 1,$$

where $p_{ij}(t)$ denotes the transition probability from state i on link n-1 to state j on link n at time t and can be expressed mathematically as the corresponding observed frequencies:

$$p_{ij}(t) = Prob\{S_n(t) = j \mid S_{n-1}(t) = i\} = \frac{X_{ij}(t)}{\sum_{k=1}^{4} X_{ik}(t)} \quad (5)$$

where $S_n(t) = j$ indicates that the state on link n at time t is j and $X_{ij}(t)$ is the number of observations produced by the same vehicles experiencing state i on link n-1 and state j on link n at time t. Note that we adapt a relatively static TPM to relax the constraint of limited real-time observations and use the historical data during a 1-month period on workdays to calibrate the TPM at each time t from 7 a.m. to 8 p.m.

Given a set of real-time observations on link n-1 at time t, the state probability vector $\pi_{n-1}(t)$, consisting of the probabilities of four states, can be estimated by

$$\pi_{n-1}(t) = \begin{bmatrix} \pi_{n-1,t}^1 & \pi_{n-1,t}^2 & \cdots & \pi_{n-1,t}^4 \end{bmatrix}$$
$$= \begin{bmatrix} \dfrac{X_1(t)}{\sum_{k=1}^{4} X_k(t)} & \dfrac{X_2(t)}{\sum_{k=1}^{4} X_k(t)} & \cdots & \dfrac{X_4(t)}{\sum_{k=1}^{4} X_k(t)} \end{bmatrix} \quad (6)$$

where $\pi_{n-1,t}^1$ is the probability of state 1 on link n-1 at time t and $X_k(t)$ is the number of observations of state k on link n-1 at time t. Finally, combined with (4)~(6), the current state probabilities $\pi_n(t)$ on link n at time t are calculated as

$$\pi_n(t) = \begin{bmatrix} \pi_{n,t}^1 & \pi_{n,t}^2 & \cdots & \pi_{n,t}^4 \end{bmatrix} = \pi_{n-1}(t) P_{n-1,n}(t) \quad (7)$$

4.3. Importance Distribution Construction. A straightforward approach for estimation of link TTD is to take advantage of only real-time observations from FCD. However, under many circumstances, it is very difficult to represent the potentially multiple states with an incomplete collection of observations due to the low penetration rate of floating cars. Consequently, the chief challenge in our methodology is to construct an importance distribution (also often referred to as the proposal distribution) properly providing a high similarity to the posterior distribution, from which a series of samples with weights are randomly generated for the heterogeneous traffic states. More importantly, the importance distribution can be modified with the real-time observations to further approximate the posterior distribution.

Here we consider an importance distribution that consists of four historical normal distributions. And each historical normal distribution for the corresponding traffic state is fitted by historical observations located in the same state boundaries at time t. Combining with the estimated state probability vector $\pi_n(t)$, the importance distribution $p(x_{n,t} \mid x_{n-1,t})$ on link n incorporating the observation $x_{n-1,t}$ on link n-1 at time t can be adjusted as

$$p(x_{n,t} \mid x_{n-1,t}) = \sum_{i=1}^{4} \pi_{n,t}^i N(x_{n,t} \mid \mu_{n,t}^i, \sigma_{n,t}^i) \quad (8)$$

where $\mu_{n,t}^i$ is the mean of historical normal distribution for state i on link n at time t and $\sigma_{n,t}^i$ is the standard deviation of historical normal distribution for state i on link n at time t.

4.4. Sampling Strategy Design. In this section, a sampling strategy is developed from the concept of the particle filter approach to generate a set of high weighted particles for representing the heterogeneous traffic states. Each particle refers to a travel time sample with a corresponding weight. Here a general particle filter approach mainly consists of the state update process showing a system evolving with time and the measurement update process showing the relationship between the system states and the system outputs at a certain time. Finally, a posterior distribution can be represented by a collection of weighted particles using the particle filter approach. More details about this approach are given in the work of Doucet et al. [22].

Similar to the particle filter approach, the proposed sampling strategy comprises three components, that is, state update, measurement update, and resampling scheme. First, the importance distribution based on the spatial TPMs is applied to produce the particles instead of using a temporal state-transition model in the state update process. Then, with the arrival of new observations, probability densities as weights are updated by the likelihood function with input of the particles and real-time observations. Note that, in Figure 2(a), each observation from FCD has an error distributed around the actual mean from RFID. More importantly, the means from FCD and RFID are quite close. If we take the means from FCD as ground truth, then it is reasonable to assume that errors between particles and the mean of observations in 15-minute interval can be described as a normal distribution. In this way, the smaller error between a particle and the mean achieves the larger weight assigned to the corresponding particle. Hence, the likelihood function is represented by a zero-mean normal distribution with a standard deviation of 50 seconds based on empirical evaluation (i.e., N(0, 50)). Finally, a resampling procedure is used to eliminate particles with low weights and concentrate on particles with large weights. Residual, stratified, and systematic resampling are three common resampling schemes. In this paper, systematic resampling is preferred to be easily implemented since it aims to minimize the Monte Carlo variation and save computational time.

To summarize, the sampling strategy is specified as follows.

Step 1 (state update). Set the number of particles $N = 100$ and generate N particles $y_{n,t}^i$ with the weights $w_{n,t}^i$ (the weight refers to the probability density of the random sample) according the importance distribution $p(x_{n,t} \mid x_{n-1,t})$ on link n at time t:

$$y_{n,t}^i \sim p(x_{n,t} \mid x_{n-1,t}), \quad i \in [1, N] \quad (9)$$

Step 2 (measurement update).

(2.1) Calculate the mean of M observations $x_{n,t}^j$ from FCD on link n at time t:

$$z_{mean}^t = \frac{\sum_{j}^{M} x_{n,t}^j}{M}, \quad j \in [1, M] \quad (10)$$

(2.2) Update the weights by likelihood function p_z:

$$w_{n,t}^i = w_{n,t}^i p_z \left(y_{n,t}^i - z_{mean}^t \right) \tag{11}$$

Step 3 (resampling). Normalize the weights to ensure that the sum of weights is equal to one. Thereafter, the systematic resampling method, similar to the roulette wheel strategy in a basic genetic algorithm, is applied to retain and replicate the particles with larger weights. For more details of the method the reader can refer to [23].

4.5. Travel Time Distribution Estimation. As mentioned in the literature review, travel times on urban links may have multimodal property which are associated with the heterogeneous traffic states. Hence, Gaussian mixture model (GMM), from which the flexible fitting structure and explanatory power can be provided, is preferred to fit multistate link TTD.

Specifically, a GMM for travel time with finite K components can be formulated as follows:

$$p (x \mid \Theta) = \sum_{k=1}^{K} \pi_k \mathrm{N} \left(x \mid \mu_k, \sigma_k^2 \right) \tag{12}$$

where π_k is the mixture coefficient for the kth normal distribution such that $\pi_k \geq 0$ and $\sum_{k=1}^{K} \pi_k = 1$. μ_k and σ_k^2 are the mean and variance for the kth normal distribution, respectively. $\Theta = \{(\pi_1, \pi_2, \ldots, \pi_K), (\mu_1, \mu_2, \ldots, \mu_K), (\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2)\}$ is the set of GMM parameters. $x = (x_1, x_2, \ldots, x_N)$ is a vector of observations. In this paper, the number of components in GMM ranges from $K = 2$ to $K = 4$ in terms of four states in total.

To estimate the GMM parameters, EM algorithm and MCMC simulation are two commonly used methods. However, the EM algorithm based on the asymptotic theory needs samples, and it may result in a local maximum problem [24]. Unlike the EM algorithm, the MCMC simulation as a golden method for Bayesian inference can provide a posterior distribution of the parameters in the proposed model, which incorporates prior knowledge with smaller number of sample size in an iterative sampling procedure [25]. Such a sampling process would repeat many times until the estimated parameters tend to converge and produces the posterior distribution.

In Bayesian inference, the posterior distribution $p(\Theta \mid x)$ for GMM can be expressed in terms of the prior distribution $f(\Theta)$ and likelihood function $L_x(x \mid \Theta)$.

$$p (\Theta \mid x) = \frac{L_x (x \mid \Theta) f (\Theta)}{f (x)} \propto L_x (x \mid \Theta) f (\Theta) \tag{13}$$

$$L_x (x \mid \Theta) = \prod_{i=1}^{N} \sum_{k=1}^{K} \pi_k \mathrm{N} \left(x_i \mid \mu_k, \sigma_k^2 \right) \tag{14}$$

$$f (\Theta) = \prod_{k=1}^{K} p \left(\mu_k \mid \sigma_k^2, \pi_k \right) p \left(\sigma_k^2 \mid \pi_k \right) p (\pi_k) \tag{15}$$

where according to the conjugacy property that describing the posterior distribution has the same parametric form as the prior distribution, the priors for distributions $\{\mu, \sigma, \pi\}$ are assumed to be as follows [24]:

$$\mu_k \mid \sigma_k^2, \pi_k \sim \mathrm{Normal} \left(b, \frac{\sigma_k^2}{\tau} \right) \tag{16}$$

$$\frac{1}{\sigma_k^2} \mid \pi_k \sim \mathrm{Gamma} (\alpha, \beta) \tag{17}$$

$$\pi_k \sim \mathrm{Dirichlet} \left(\frac{1}{K}, \frac{1}{K} \ldots, \frac{1}{K} \right) \tag{18}$$

where parameters $\{b, \tau, \alpha, \beta\}$ are taken as known 'hyperparameters'. In the initialization of the MCMC simulation, the hyperparameters can be set to $\{b = 150, \tau = 0.1, \alpha = 1, \beta = 2\}$. Finally, the pseudocode of the proposed MCMC method using Gibbs sampling algorithm is presented in Table 1.

Another important point is that the number of components in GMM is unknown. An information criterion, known as Bayesian Inference Criterion (BIC) for the goodness of fit assessment, is chosen to find the optimal number of components [26]. In use of the BIC, a smaller value of BIC indicates a better fitting model. Specifically, the BIC taking into account the log likelihood, the number of estimated parameters d, and sample size N is defined as

$$BIC (K) = -2 \ln L_x (x \mid \Theta) + d \ln N \tag{19}$$

5. Case Study and Verification

5.1. Multistate Travel Time Histograms. Before utilizing the proposed method for estimation of multistate link TTD, the simple histogram can exhibit intuitive information on the TTD. Figure 5 shows two or more clusters in travel time histograms for two different traffic conditions.

During off-peak period, Figure 5(a) gives a distinctly bimodal pattern for the observations from RFID, and most of the observations fall into state 1 and state 3. Conversely, during peak period, as shown in Figure 5(b), state 4 gathering over 80% of the observations plays a dominant role in the congested situation. In terms of the FCD, one (Figure 5(a)) is very small number of observations in state 1 and state 2 whereas there are no observations in state 3, and the other (Figure 5(b)) is the observations largely distributed beyond the two peaks in state 3 and state 4, which represent the major clusters of travel times.

Overall, this simple example implies that the issues of sparse and incomplete observations from FCD may result in irregular, unsmooth TTDs. To improve the estimation of link TTDs, it is suggested that historical information on prior travel time patterns is strongly needed.

5.2. Estimation of Link TTD. For each number of components, the sampling procedure was run in a total of 20,000 iterations and the first 10,000 iterations were discarded as burn-in, while the remaining samples were used for inferencing the posterior distribution. Convergence of the

TABLE 1: Gibbs sampling algorithm for a K component GMM combining with BIC.

Input: N particles (i.e. $x_1, x_2, \ldots, x_N$) with large weights selected by sampling strategy, K

Output: parameters $\Theta = \{(\pi_1, \pi_2, \ldots, \pi_K), (\mu_1, \mu_2, \ldots, \mu_K), (\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2)\}$, $BIC(K)$

Step 1. Initialization

Determine hyperparameters: b, τ, α, β

Set iteration $r = 0$

Set $e_k^{(r)} = 1/K$, draw $\pi_k^{(r)} \sim \text{Dirichlet}(e_1^{(r)}, \cdots, e_K^{(r)})$

Draw $\mu_k^{(r)} \sim \text{Normal}(b_k^{(r)}, \sigma_k^{2(r)}/\tau_k^{(r)})$

Draw $1/\sigma_k^{2(r)} \sim \text{Gamma}(\alpha_k^{(r)}, \beta_k^{(r)})$

Step 2. Gibbs sampling

For $r = 1$ to R

Update the mixing coefficients $\pi_k^{(r)}$:

Draw $\pi_k^{(r)} \sim \text{Dirichlet}(e_1^{(r)} + E_1^{(r)}, \cdots, e_K^{(r)} + E_K^{(r)})$, where, $E_K^{(r)}$ is the effective number of particles assigned to component K, and can be calculated by

$$E_k^{(r)} = \sum_{i=1}^{N} \frac{\pi_k^{(r-1)} N\left(x_i \mid \mu_k^{(r-1)}, \sigma_k^{2(r-1)}\right)}{\sum_{k=1}^{K} \pi_k^{(r-1)} N\left(x_i \mid \mu_k^{(r-1)}, \sigma_k^{2(r-1)}\right)}$$

For $k = 1$ to K

Update the variance $\sigma_k^{2(r)}$:

Draw $1/\sigma_k^{2(r)} \sim \text{Gamma}(\alpha_k^{(r)}, \beta_k^{(r)})$, where

$$\alpha_k^{(r)} = \alpha_k^{(r-1)} + \frac{E_k^{(r)}}{2}$$

$$\beta_k^{(r)} = \beta_k^{(r-1)} + \frac{1}{2}\left(\sum_{i=1}^{N}(x_i - \bar{x}_k)^2 + \frac{\left(b_k^{(r-1)} - \bar{x}_k\right)^2 \cdot E_k^{(r)} \cdot \tau_k^{(r-1)}}{E_k^{(r)} + \tau_k^{(r-1)}}\right)$$

Update the mean $\mu_k^{(r)}$:

Draw $\mu_k^{(r)} \sim \text{Normal}(b_k^{(r)}, \sigma_k^{(r)}/\tau_k^{(r)})$, where

$$b_k^{(r)} = \frac{\tau_k^{(r-1)} \cdot b_k^{(r-1)} + E_k^{(r)} \bar{x}_k}{\tau_k^{(r-1)} + N_k^{(r)}}$$

$$\tau_k^{(r)} = \tau_k^{(r-1)} + E_k^{(r)}$$

End for

End for

Step 3. BIC calculation

Calculate the BIC for a K component GMM using Eq. (19)

Note: we first execute the sampling with increasing number of components from $K = 2$ to $K = 4$.

(a) a.m. off-peak period (07:00-07:15)

□ RFID observations (52 points)
■ FCD observations (5 points)

(b) a.m. peak period (08:30-08:45)

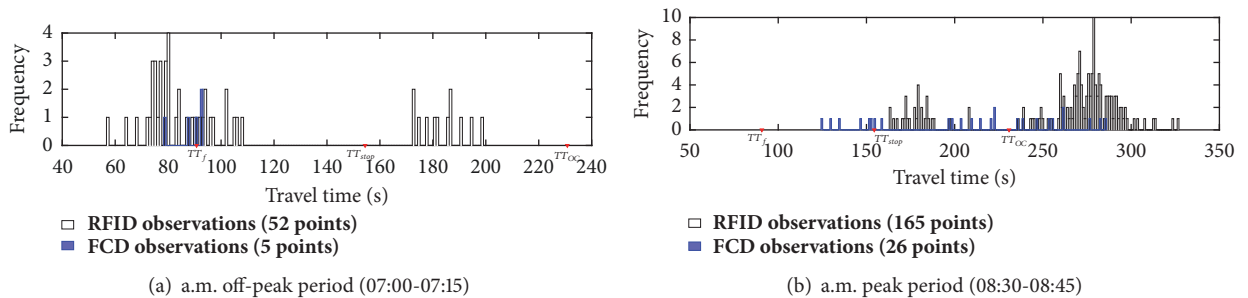□ RFID observations (165 points)
■ FCD observations (26 points)

FIGURE 5: Travel time histograms of study link on Dec. 1, 2011 (Thursday).

estimated parameters was monitored by the trace plots of the generated sample values, the model log likelihood, and the autocorrelations [26].

To evaluate the performance of the proposed method, the Hellinger distance (HD) and Kolmogorov-Smirnov (KS) test are employed to assess the similarity between the estimated distribution from FCD and the empirical (actual) one from RFID. For discrete distributions, the HD between the estimated distribution TTD_{est} and the empirical one TTD_{emp} is defined as follows [26]:

$$H\left(\text{TTD}_{est}, \text{TTD}_{emp}\right)$$

$$= \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1} \left(\sqrt{f_{est}(x_i)} - \sqrt{f_{emp}(x_i)} \right)^2} \qquad (20)$$

where $f_{est}(i)$ and $f_{emp}(i)$ are the estimated and observed probabilities for an observation x_i from RFID. Besides, the HD satisfies two requirements in (21).

$$H\left(\text{TTD}_{est}, \text{TTD}_{emp}\right) = H\left(\text{TTD}_{emp}, \text{TTD}_{est}\right),$$

$$0 \le H\left(\text{TTD}_{est}, \text{TTD}_{emp}\right) \le 1 \qquad (21)$$

It is indicated that HD has the nature of symmetry and nonnegativity. Furthermore, if $H(\text{TTD}_{est}, \text{TTD}_{emp}) = 0$, the two distributions will be close to each other, whereas $H(\text{TTD}_{est}, \text{TTD}_{emp}) = 1$ means that there exists the most difference between distributions.

The KS test is a common nonparametric test used to test if the estimated distribution passes the null hypothesis that the observations are distributed according to the estimated distribution. A larger p value provided from the test indicates a better goodness of fit. The estimated distribution is rejected when p value is smaller than 0.05 at a 95% confidence level.

Figure 6 shows a comparison of the estimated TTD and the empirical one. It indicates that if the importance distribution is not accurate enough as shown in Figure 6(a), the observations from FCD imply that there is no data in between the two peaks, and then Gaussian likelihood function tends to pull the posterior distribution in the right direction. Besides, the results of link TTD estimation show the advantage of the proposed approach with few data in terms of the calibration of TPMs from historical travel times.

If the importance distribution is quite accurate, as shown in Figure 6(b), however, the posterior distribution is close to the empirical one, the Bayesian posterior distribution says that some data with high probabilities are missing in the first peak (around 185 s). If we observe the travel time histogram at a.m. peak periods in Figure 5(b), eight observations from FCD located in between two major components from RFID, there should not be such an obvious gap between state 3 and state 4 under congested situation. Besides, the observations nearby a mean of 210 seconds in the Gaussian likelihood function will be assigned larger weights. Altogether, with the Bayesian update, the posterior distribution that incorporates information not only from the prior travel time pattern, but also from the real-time data, may provide a more convincing result.

Table 2 summarizes the model parameters and performance results. It can be seen that the GMM with three components provides a superior fit for empirical data during peak hours. This could be explained by the underlying situations that present two kinds of state combinations, that is, state 1 to state 3, and state 2 to state 4, while the larger weight implies the corresponding state in a state combination may have a leading impact on the whole link traffic conditions. These findings assist with insight into the heterogeneity in travel times. Moreover, the performance results show that of all 52 periods used for the KS test, only nine of them fail to pass the

KS test, which help confirm the statistical superiority of the proposed method.

5.3. Effect of Floating Car Percentages on the Proposed Approach. To examine the robust of the proposed method, six different percentages of floating cars, i.e., floating cars% = 0%, 5%, 10%, 20%, 50%, and 100%, are conducted as an input of estimating the link TTDs. For each percentage of floating cars, the sampling procedure was run 100 times to randomly select the corresponding samples from observations and calculate the average values of HD and percentages of passing KS test in each 15-min interval. In particular, if the number of samples is zero, we use an alternative method with the steady-state probabilities and the means of historical travel times instead of the estimated state probabilities in (7) and the means of real-time observations in likelihood function.

According to the irreducible and ergodic properties of the Markov chain, the steady-state TPM $\Pi_n(t)$ on link n at time t can be defined as follows [27]:

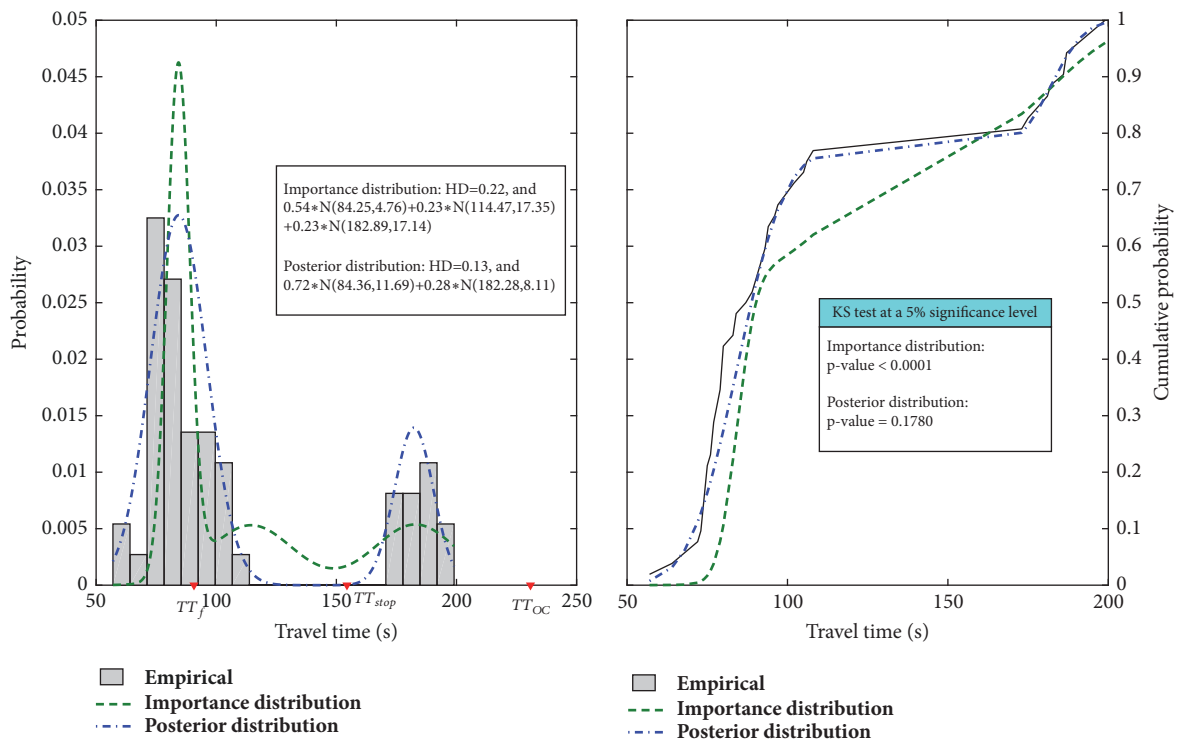$$\Pi_n(t) = \lim_{r \to \infty} P^r_{n,n-1}(t) \qquad (22)$$

where r is the number of steps. When r is big enough, each row of $\Pi_n(t)$ will all reach steady-state probabilities regardless of whatever the state probabilities on link n-1 are. Figure 7 shows the results of average HD and KS test for different percentages of floating cars during time of day.

It can be observed that, with the percentage of floating cars decreasing, the main trend of the average HD increases; on the contrary, the percentage of passing KS test in general decreases. Especially when the number of samples is zero, the performance of the alternative method (i.e., the average HD largely varying between 0.1 and 0.3 while the percentage of passing KS test mainly ranging from 70% to 90%) is superior to the proposed method with sample size less than five, which mainly generates the worst result. This is likely due to the random sampling resulting in higher travel time variability for estimation input, whereas the alternative method only depending on the historical travel time information is not sensitive to the input data.
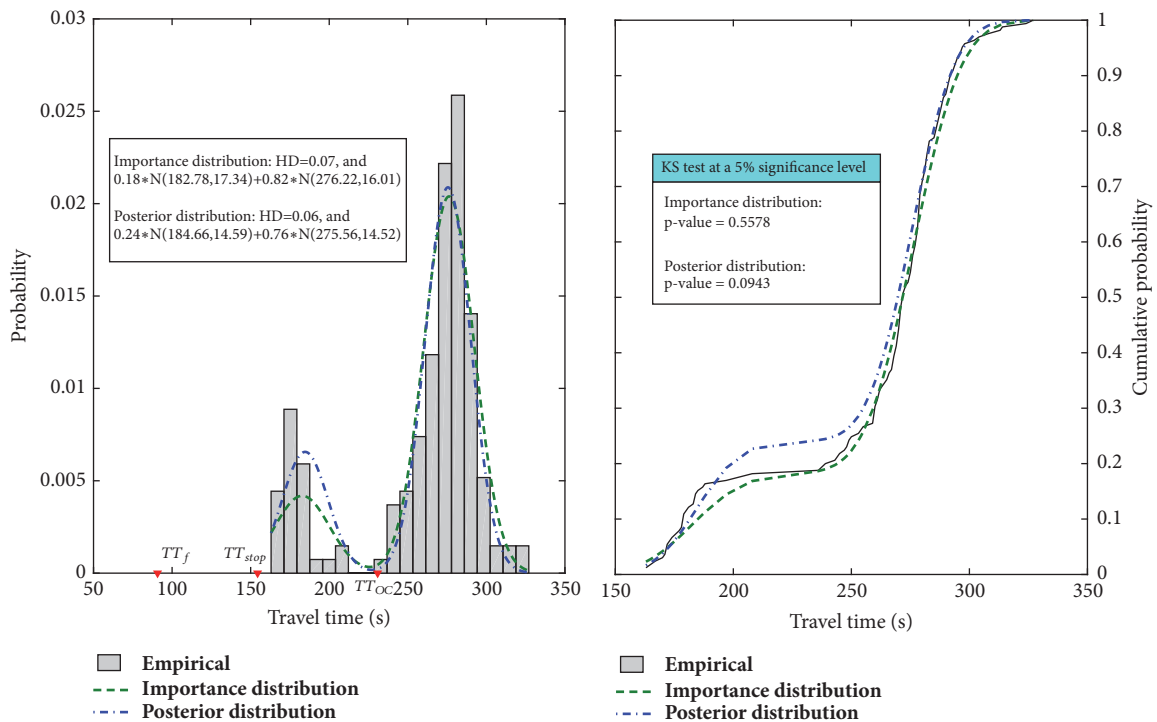
Overall, the results confirm that the proposed method still performs well with small sample size. For the worst condition, i.e., floating cars% = 0%, the alternative method is preferred in terms of the estimation performance.

6. Conclusions

This paper presents a Bayesian approach based on particle filter framework to estimate link TTDs using the limited observations and historical data set from FCD. The main objective of the approach is to produce the multistate TTDs with the low penetrate rate of floating cars. Unlike previous studies that require a complete data in modeling the multistate TTDs, the proposed approach uses less observations to update the importance distribution. Concisely, the importance distribution is regarded as prior knowledge and provides past travel time patterns. On the other hand, limited observations are integrated into the particle filter framework and make the estimated TTDs further close to the empirical

Importance distribution: HD=0.22, and
0.54*N(84.25,4.76)+0.23*N(114.47,17.35)
+0.23*N(182.89,17.14)

Posterior distribution: HD=0.13, and
0.72*N(84.36,11.69)+0.28*N(182.28,8.11)

KS test at a 5% significance level

Importance distribution:
p-value < 0.0001

Posterior distribution:
p-value = 0.1780

Empirical
Importance distribution
Posterior distribution

(a)

Importance distribution: HD=0.07, and
0.18*N(182.78,17.34)+0.82*N(276.22,16.01)

Posterior distribution: HD=0.06, and
0.24*N(184.66,14.59)+0.76*N(275.56,14.52)

KS test at a 5% significance level

Importance distribution:
p-value = 0.5578

Posterior distribution:
p-value = 0.0943

Empirical
Importance distribution
Posterior distribution

(b)

FIGURE 6: Comparison of the estimated TTD from FCD and the empirical TTD from RFID for the periods of (a) 07:00-07:15 a.m. and (b) 08:30-08:45 a.m.

TABLE 2: Link TTD parameters, HD, and KS tests. This table is reproduced from a shorter version of this work presented at the 97th Annual Meeting of the Transportation Research Board [5].

| Time | Sample size (N) | Estimated parameters in posterior distribution | | | HD | h^a | KS test |
		Weight (π)	Mean (μ)	Standard deviation (σ)			p value
			a.m. Peak 7:00 - 9:00				
7:00-7:15	5	(0.72, 0.28)	(84.36, 182.28)	(11.69, 8.11)	0.13	0	0.1780
7:15-7:30	25	(0.68, 0.22, 0.1)	(84.29, 114.66, 181.87)	(10.01, 10.03, 10.055)	0.09	0	0.0865
7:30-7:45	29	(0.48, 0.2, 0.32)	(83.87, 115.81, 197.72)	(14.11, 19.11, 14.02)	0.08	1	0.0305
7:45-8:00	22	(0.46, 0.42, 0.12)	(96.48, 136.53, 203.14)	(10.23, 9.53, 9.5)	0.17	0	0.0797
8:00-8:15	28	(0.71, 0.17, 0.12)	(132.51, 183.23, 248.51)	(20.13, 20.01, 20.36)	0.08	0	0.2145
8:15-8:30	27	(0.53, 0.05, 0.42)	(143.34, 183.5, 269.59)	(20.21, 6.67, 10.17)	0.24	0	0.7334
8:30-8:45	26	(0.24, 0.76)	(184.66, 275.56)	(14.59, 14.52)	0.06	0	0.0943
8:45-9:00	21	(0.21, 0.1, 0.69)	(182.52, 222.52, 275.02)	(8.08, 12.01, 13.15)	0.11	1	0.0325
			p.m. Peak 17:00 - 19:00				
17:00-17:15	17	(0.59, 0.41)	(104.72, 205.44)	(19.82, 13.07)	0.16	0	0.4162
17:15-17:30	24	(0.29, 0.56, 0.16)	(105.83, 207.03, 246.86)	(17.38, 9.23, 25)	0.09	0	0.4507
17:30-17:45	14	(0.24, 0.58, 0.18)	(123.7, 211.82, 251.41)	(9.55, 12.1, 32.82)	0.11	1	0.0022
17:45-18:00	16	(0.14, 0.48, 0.38)	(147.92, 218.88, 237.36)	(20.12, 6.93, 10.01)	0.14	0	0.1100
18:00-18:15	20	(0.35, 0.65)	(117.46, 216.85)	(20, 20.18)	0.06	0	0.0928
18:15-18:30	27	(0.43, 0.29, 0.28)	(108, 191.79, 206.68)	(21.3, 20, 20.39)	0.12	0	0.1824
18:30-18:45	10	(0.45, 0.13, 0.42)	(90.85, 129.37, 196.79)	(10.13, 20.01, 10.01)	0.28	0	0.7900
18:45-19:00	12	(0.7, 0.08, 0.22)	(95.04, 130.72, 193.38)	(12.57, 8.99, 17.53)	0.09	0	0.6123
			Performance summary during time of day				

Time of day 7:00-20:00	Average HD 0.121	Min HD 0.024	Max HD 0.355	Percentages of passing KS test (%) 83

^{a}h = 0 indicates that the estimated distribution passes the KS test at a 5% significance level.

(a) p.m. period of 17:15 to 17:30

(b) Comparison of the values of average HD
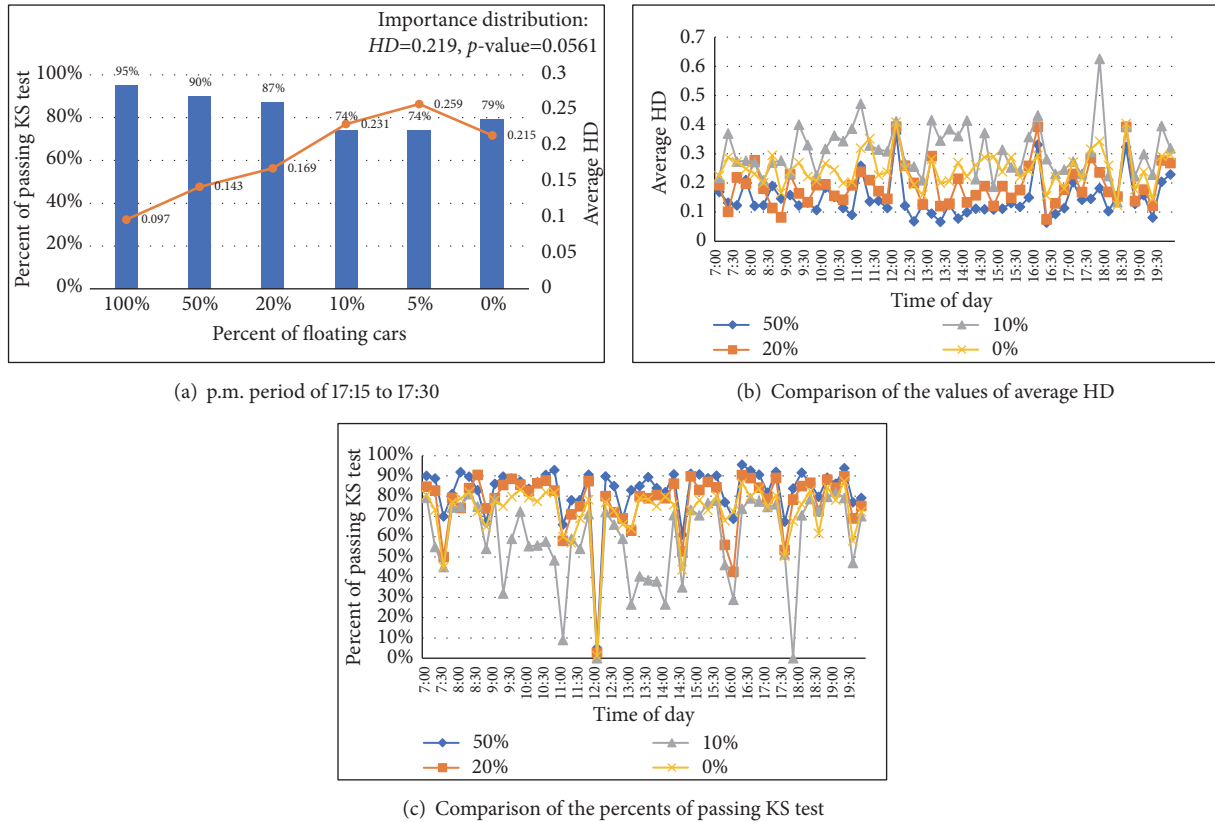

(c) Comparison of the percents of passing KS test

FIGURE 7: The results comparisons for different percentages of floating cars. This figure is reproduced from a shorter version of this work presented at the 97th Annual Meeting of the Transportation Research Board [5].

distributions by using a sampling strategy. Finally, the result is a posterior distribution incorporating historical trends and the real-time travel time information as well.

The results of the case study demonstrate that the proposed approach produces the posterior distribution agreeing well with the empirical distribution. Besides, with different percentages of floating cars, our proposed approach performs equally well, even under condition with no data available. This suggests that, when the number of floating car sample size is less than five, the approach with the steady-state probabilities helps provide a promising result.

Nevertheless, this paper has some limitations. First, we used the historical observations as an a priori estimation. Information from estimated TTD in previous period should be integrated into the particle filter framework to update the parameters of the current importance distribution and provide a better a priori estimation. Second, the spatial TPMs of the Markov chain did not consider the temporal correlation between link travel times. A Markov process with a short memory taking into account the effect of travel time states at the preceding time interval on transition intensities may lead to better TPMs estimates. Finally, a density-based clustering method such as DBSCAN clustering or other DBSCAN variants should be developed to cluster link travel times for representing heterogeneous traffic states so that the ability of the proposed approach could be generalized to a large range of applications from urban signalized arterials to freeway

continuous traffic instead of using traffic engineering rules. In addition, the effectiveness of the proposed approach still needs a further verification by conducting more case studies in other road links with different sources of trajectory data (e.g., smartphone and the ride service platform of Uber or DiDi Chuxing). These will be our future works. Last but not least, we should state that the work from this paper is an extended version of the earlier work presented at the 97th Annual Meeting of the TRB, Washington, DC, January 2018 [5]. For more details, it can be found in [5].

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] Y.-X. Ji and H. Zhang, "Travel Time Distributions on Urban Streets: Their Estimation with a Hierarchical Bayesian Mixture

Model and Application to Traffic Analysis using High - Resolution Bus Probe Data," in *93th Annual Meeting of the Transportation Research Board*, D.C, Washington, 2014.

[2] M. A. P. Taylor and S. Somenahalli, "Travel time reliability and the bimodal travel time distribution for an arterial road," *Road & Transport Research*, vol. 19, no. 4, pp. 37–50, 2010.

[3] F. Yang, M.-P. Yun, and X.-G. Yang, "Travel time distribution under interrupted flow and application to travel time reliability," *Transportation Research Record*, vol. 2466, pp. 114–124, 2014.

[4] M.-P. Yun, W.-W. Qin, Y. Liu, and J. Yu, "Analysis of Required Minimum Sample Size of Floating Cars for Estimating Urban Road Link Travel Time Considering Bimodal Distribution and Estimation Error," in *96th Annual Meeting of the Transportation Research Board*, D.C, Washington, 2017.

[5] W.-W. Qin and M.-P. Yun, "Estimation of Urban Link Travel Time Distribution Using Markov Chains and Bayesian Approaches," in *97th Annual Meeting of the Transportation Research Board*, D.C, Washington, 2018.

[6] D. Dailey and F. Cathey, "AVL-Equipped Vehicles as Traffic Probe Sensors," Technical Report No. WA-RD 534.1, Washington State Transportation Center (TRAC), 2002.

[7] Y.-H. Feng, J. Hourdos, and G. A. Davis, "Probe vehicle based real-time traffic monitoring on urban roadways," *Transportation Research Part C: Emerging Technologies*, vol. 40, pp. 160–178, 2014.

[8] T. Hunter, R. Herring, and A. Bayen, "Path and Travel Time Inference from GPS Probe Vehicle Data," in *Proceedings of the In Neural Information Processing Systems Foundation (NIPS) Conference*, Vancouver, Canada, 2009.

[9] R. Herring, A. Hofleitner, P. Abbeel, and A. Bayen, "Estimating arterial traffic conditions using sparse probe data," in *Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems, ITSC 2010*, pp. 929–936, Portugal, September 2010.

[10] B. S. Westgate, D. B. Woodard, D. . Matteson, and S. G. Henderson, "Travel time estimation for ambulances using Bayesian data augmentation," *The Annals of Applied Statistics*, vol. 7, no. 2, pp. 1139–1161, 2013.

[11] F. Guo, Q. Li, and H. Rakha, "Multistate travel time reliability models with skewed component distributions," *Transportation Research Record*, vol. 2315, pp. 47–53, 2012.

[12] S. Park, H. Rakha, and F. Guo, "Multi-state travel time reliability model: Impact of incidents on travel time reliability," in *Proceedings of the 14th IEEE International Intelligent Transportation Systems Conference, ITSC 2011*, pp. 2106–2111, USA, October 2011.

[13] A. Hofleitner, R. Herring, and A. Bayen, "Arterial travel time forecast with streaming data: A hybrid approach of flow modeling and machine learning," *Transportation Research Part B: Methodological*, vol. 46, no. 9, pp. 1097–1122, 2012.

[14] Z.-G. Li, C. Cai, A. K. Menon, Y. Xu, and F. Chen, "Estimation of link speed distribution from probe vehicle data," *Transportation Research Record*, vol. 2595, pp. 98–107, 2016.

[15] M. Ramezani and N. Geroliminis, "On the estimation of arterial route travel time distribution with Markov chains," *Transportation Research Part B: Methodological*, vol. 46, no. 10, pp. 1576–1590, 2012.

[16] Y. Lou, C. Zhang, Y. Zheng, X. Xie, W. Wang, and Y. Huang, "Map-matching for low-sampling-rate GPS trajectories," in *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 352–361, ACM, November 2009.

[17] I. Sanaullah, M. Quddus, and M. Enoch, "Developing travel time estimation methods using sparse GPS data," *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, vol. 20, no. 6, pp. 532–544, 2016.

[18] Z.-L. Ma, L. Ferreira, M. Mesbah, and S. Zhu, "Modeling distributions of travel time variability for bus operations," *Journal of Advanced Transportation*, vol. 50, no. 1, pp. 6–24, 2016.

[19] D. J. Sun, X. Liu, A. Ni, and C. Peng, "Traffic congestion evaluation method for urban arterials: Case study of Changzhou, China," *Transportation Research Record*, vol. 2461, pp. 9–15, 2014.

[20] F. Yang, *Research on urban network traffic flow evolution based on experimental traffic theory and mathematical analysis [Ph.D. thesis]*, College of transportation engineering, Tongji University, China, 2015.

[21] Transportation Research Board, Transportation Research Board. Highway Capacity Manual.

[22] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*, Springer, New York, NY, USA, 2001.

[23] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[24] B.-J. Park, Y. Zhang, and D. Lord, "Bayesian mixture modeling approach to account for heterogeneity in speed data," *Transportation Research Part B: Methodological*, vol. 44, no. 5, pp. 662–673, 2010.

[25] E. Kidando, R. Moses, E. E. Ozguven, and T. Sando, "Bayesian nonparametric model for estimating multistate travel time distribution," *Journal of Advanced Transportation*, vol. 2017, Article ID 5069824, 9 pages, 2017.

[26] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. van der Linde, "Bayesian measures of model complexity and fit," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 64, no. 4, pp. 583–639, 2002.

[27] J. Yeon, L. Elefteriadou, and S. Lawphongpanich, "Travel time estimation on a freeway using Discrete Time Markov Chains," *Transportation Research Part B: Methodological*, vol. 42, no. 4, pp. 325–338, 2008.