Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 1793060, 16 pages
https://doi.org/10.1155/2022/1793060

Research Article

Urban Traffic State Estimation with Online Car-Hailing Data: A Dynamic Tensor-Based Bayesian Probabilistic Decomposition Approach

Wenqi Lu,[1,2] **Ziwei Yi**,[1,2] **Dongyu Luo**,[3] **Yikang Rui**,[1,2] **Bin Ran**,[1,2] **Jianqing Wu**,[4] **and Tao Li**[5]

[1]*School of Transportation, Southeast University, Nanjing, Jiangsu 211189, China*
[2]*Joint Research Institute on Internet of Mobility, Southeast University and University of Wisconsin-Madison, Southeast University, Nanjing 211189, China*
[3]*Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 10044, China*
[4]*School of Qilu Transportation, Shandong University, Jinan 250002, China*
[5]*Shandong Hi-speed Construction Management Group Co., Ltd., Jinan 250000, China*

Correspondence should be addressed to Yikang Rui; 101012189@seu.edu.cn

Received 2 January 2022; Accepted 4 March 2022; Published 26 April 2022

Academic Editor: Sang-Bing Tsai

Timely and precise traffic state estimation of urban roads is significant for urban traffic management and operation. However, most of the advanced studies focus on building complex deep learning structures to learn the spatiotemporal feature of the urban traffic flow, ignoring improving the efficiency of the traffic state estimation. Considering the benefit of the tensor decomposition, we present a novel urban traffic state estimation based on dynamic tensor and Bayesian probabilistic decomposition. Firstly, the real-time traffic speed data are organized in the form of a dynamic tensor which contains the spatiotemporal characteristics of the traffic state. Then, a dynamic tensor Bayesian probabilistic decomposition (DTBPD) approach is built by decomposing the dynamic tensor into the outer product of several vectors. Afterward, the Gibbs sampling method is introduced to calibrate the parameters of the DTBPD models. Finally, the real-world traffic speeds data extracted from online car-hailing trajectories are employed to validate the model performance. Experimental results indicate that the proposed model can greatly reduce computational time while maintaining relatively high accuracy. Meanwhile, the DTBPD model outperforms the state-of-the-art models in terms of both single-step-ahead and multistep-ahead traffic state estimation.

1. Introduction

With the rapid advancement of urbanization, urban traffic congestion has become a critical problem in the construction of smart cities [1]. It may take place when the traffic demand is beyond the capacity of the road network, resulting in more traffic accidents, severe air pollution, and increased fuel consumption [2]. As a significant part of the intelligent transportation system and automated vehicle technology [3], accurate estimation of the urban road network traffic state in the near future is very important for traffic management and planning in reducing traffic congestion. By providing accurate and comprehensive network state estimation information, human-driven vehicles and connected automated vehicles (CAV) can optimize travel paths and generate lane-changing recommendations [4]. Fleet management companies are capable of operating the dispatch system more efficiently [5]. Traffic management agencies can update the control plan in real time to improve control performance. Besides, the traffic policy-making department can analyze emerging needs and evaluate the policy impact [6, 7].

During these years, many approaches have been proposed for the estimation traffic state of the urban roads including statistical models [8], artificial-based intelligent models [9], and deep learning-based models [10]. From the perspective of the research objects, the macroscopic parameters of the lane sections, intersection, link sections, and road networks, such as travel speed, volume, and density, are usually considered as the estimated targets [11].

However, the challenges of the urban traffic state estimation still exist as follows. (1) The collection of urban traffic data may still rely on stationary detectors or cameras, which is expensive and inflexible. It is not usually a realistic choice to deploy sufficient detectors on low-level roads such as secondary roads and access roads. The traffic states of some low-level roads are always neglected though they may have a significant influence on the future state of the urban network [11]. (2) Though the deep learning-based models are capable of learning the spatiotemporal feature of urban traffic state, the performance of these methods is partly influenced by the quantity and quality of the research data. Achieving a well-trained deep learning model is time-consuming, and it is difficult to update the parameters of the trained models in a short time.

Nowadays, the massive trajectory data of various car-hailing companies have become a popular and alternative source for analyzing traffic operation and traffic emissions [12–14], and it also provides the basis for the urban traffic state estimation since the trajectory data are high-precision, high-resolution, and widely distributed. In addition, the tensor-based approaches turn out to be efficient for dealing with traffic data due to their ability to capture the multimode relevance of data [15–18], and it provides a new perspective on modeling.

Inspired by the rise of ride-hailing trajectory data and the benefit of tensor mode, this paper proposes a traffic state estimation model based on dynamic tensor Bayesian probabilistic decomposition (DTBPD). Firstly, the concept of the dynamic tensor is introduced to form the urban spatiotemporal data into the tensor pattern. Then a higher-order Bayesian matrix factorization model is established based on CANDECOMP/PARAFAC (CP) decomposition. Meanwhile, the Gibbs sampling algorithm is established to perform inference on the element of the Bayesian decomposition model. Finally, we utilize the urban traffic speed data captured from online car-hailing trajectory data provided by the Didi company to validate the feasibility of the DTBPD approach under the scenarios of different types of roads. As far as the authors know, this is the first time that a dynamic tensor-based probabilistic paradigm has been employed for urban traffic state estimation.

The main contributions of this paper can be summarized as follows:

(1) The traffic speeds of the urban network are organized in the form of the dynamic tensor pattern with consideration of the spatiotemporal features for the first time.

(2) We propose a dynamic tensor Bayesian probability decomposition structure that is capable of modeling the dynamic tensor in real time with highly spatiotemporal data and provide a Gibbs sampling algorithm for estimating model parameters.

(3) The traffic speeds of the urban network are extracted from the online car-hailing trajectory data of Didi company and used for validating the efficiency of the DTBPD model under multiple scenarios. Experimental results demonstrate that the DTBPD model is not only accurate but also time-saving during the calibration process.

In the next section, a general overview of related work on urban traffic estimation is provided. Section 3 gives the basic definition of tensor decomposition. The methodology of the DTBPD model is introduced in Section 4. In Section 5, trajectory data description and preprocessing with data analysis are provided. The results and discussion of the experiments are introduced in Section 6. Some conclusions and future work are discussed in Section 7.

2. Related Work

The methods for estimating traffic state usually can be categorized into two classes including the parametric methods and nonparametric methods.

The common parametric methods include the statistical method, time-series method, and Kalman filter method [11]. In the time series method, the most common method is the Autoregressive integrated moving average model (ARIMA) and Ahmed and Cook [19] applied it to the traffic state estimation. Since ARIMA is only applicable to several scenarios with continuous and sufficient historical data, many scholars extended the ARIMA such as seasonal ARIMA (SARIMA) [20]. Another typical nonparametric approach is Kalman filter-based methods. Using the information of CAVs, Emami et al. [21] gave a faded memory Kalman filter to forecast the short-term flow at urban arterials. However, the parametric methods normally have fixed structures and are unable to fully learn the spatiotemporal characteristics of traffic flow.

With the development of mobile communication networks and big data, the nonparametric methods represented by machine learning models emerged to deal with the above issues [22–24]. For example, Castro-Neto et al. [25] built an online support vector machine model considering normal conditions and abnormal conditions such as accidents. Sun et al. [26] proposed a dynamic process K-nearest Neighbor method, which enables KNN parameters to be self-adjusted and robust without the need to predefine models or train parameters.

In addition, the excellent performance of deep learning (DL) in processing massive data has also attracted extensive attention, and it has been widely used in computer vision, semantic recognition, and automatic driving. In recent years, DL-based models [27–31] have been introduced into traffic flow estimation. Lv et al. [32] proposed a deep architecture model considering the temporal and spatial characteristics of the traffic flow by employing the autoencoder. To solve the long-term dependence in the

prediction process, some variants of recurrent neural networks (RNN), such as long short-term memory (LSTM) neural network [33], Gated recurrent unit (GRU) neural network [34], and Transformer [35] were proposed and applied to learning the temporal characteristics for the traffic state estimation. Ma et al. [36] represented traffic flow data as images and used a convolutional neural network-based architecture to estimate the traffic state in a large network. Furthermore, some DL-based combination models [37–40] have been proposed to improve the ability to learn spatiotemporal features and the robustness of the estimation. Wu et al. [41] proposed a traffic flow prediction model with a hybrid DL neural network, where convolutional neural networks (CNN) and RNN were used to extract spatiotemporal features, respectively, and an attention mechanism was introduced to determine the importance of historical data. Li et al. [42] employed a deep fusion model combining stacked restricted Boltzmann machines to forecast the accident duration. To capture the spatial and temporal dependences simultaneously, Zhao et al. [43] constructed a novel neural network-based traffic forecasting method named T-GCN, which combines the graph convolutional network (GCN) [44] and the GRU for traffic forecasting. Although the DL-based traffic state estimation approaches can fully mine the potential characteristics of the traffic flow, DL-based traffic state estimation approaches take a long time to train or calibrate the models, resulting in inconvenient and inflexible updates of the parameters [45]. Meanwhile, tuning the parameters of a DL-models reasonably is a very difficult task and highly depends on experience, which prevents these models from being used in practice [29].

To tackle the above problem, tensor-based methods [18, 46–48] have become new solutions for traffic estimation issues due to their efficient architecture to mine temporal and spatial correlations of the traffic flow. Ran et al. [49] built a tucker decomposition-based imputation method for the traffic state estimation of the floating car system. Further, Ran et al. [50] provided a high accuracy low-rank tensor completion (HaLRTC) algorithm for estimating missing traffic speed data, and it can address the extreme case where the data of a long period of one or several weeks are completely missing [51]. However, these models rely on trace norm minimization to find a low-rank approximation to the original incomplete tensor, and they are often prone to overfitting when the missing rate is large. To address this issue, Bayesian inference methods such as the Markov chain Monte Carlo algorithm and variational inference have been designed and used for tensor decomposition [52]. For instance, Tang et al. [53] constructed a tensor decomposition method to estimate traffic flow parameters of signalized intersections with collected vehicle trajectories. Using the large-scale and sparse GPS data generated by taxicabs, Tang et al. [54] established a novel tensor-based Bayesian probabilistic model for estimating the travel time of the road links. To efficiently capture the spatiotemporal characteristics of traffic flow with tensor pattern and improve the performance of the tensor decomposition approach, we combine the concept of the dynamic tensor with the Bayesian probabilistic decomposition (DTBPD) for the first time. By establishing a Gibbs sampling algorithm to estimate the proposed model constantly and dynamically, the future traffic state data in dynamic tensor can be generated efficiently and stably.

3. Tensor Basis

3.1. Dynamic Tensor. Tensors known as multidimensional arrays are higher-order generalizations of vectors and matrices. The d^{th}-order tensor is donated as $\mathbf{X} \in \mathbb{R}^{I_1 \times I_2 \times \ldots I_d}$. Each dimension of a multidimensional array is called a mode. A sequence of d_{th}-order tensors $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \ldots, \mathbf{X}_T$ with $\mathbf{X}_t \in \mathbb{R}^{I_1 \times I_2 \times \ldots I_{\text{d}}} (1 \le t \le T)$ can be donated as a $(d+1)^{\text{th}}$-order tensor steam where T is the maximum index of the time intervals. Figure 1 indicates an example of the tensor stream of the d_{th}-order $\mathbf{X}_t \in \mathbb{R}^{I_1 \times I_2 \times \ldots I_{\text{d}}}$.

In the task of making traffic speed estimation of urban traffic state, only spatiotemporal characteristics of the tensors from several previous time intervals in the tensor stream are necessary. Therefore, a sliding time window H at time interval t is employed to localize the tensor stream into a small one, which can be written as $\widetilde{X}_{\text{t}} = \{\mathbf{X}_{t-H+1}, \mathbf{X}_{t-H+2}, \ldots, \mathbf{X}_t\}$. Figure 2 reveals the tensor stream separated by employing a sliding time window.

3.2. CANDECOMP/PARAFAC Decomposition. Let $\widetilde{X} \in \mathbb{R}^{I_1 \times I_2 \times \ldots I_d}$ be the dynamic tensor containing the observed historical values and missing values to be forecasted, where I_k is the dimension along the kth way $k \in \{1, 2, \ldots, d\}$. $x_{i_1, i_2, \ldots, i_d} (1 \le i_k \le I_k)$ is a value of an entry in $\widetilde{X}$. The idea of CP decomposition is to approximate $\widetilde{X}$ by calculating the sum of R rank-one component tensors as follows:

$$\widetilde{X} = \sum_{r=1}^{R} \mathbf{a}_r^1 \circ \mathbf{a}_r^2 \circ \cdots \circ \mathbf{a}_r^d, \tag{1}$$

where $\circ$ represent the outer product of two vectors. $\mathbf{a}_r^k \in \mathbb{R}^{I_k}$ is the r^{th} column vector of the decomposed factor matrix $A \in \mathbb{R}^{I_k \times R}$ and $\mathbf{a}_r^1 \circ \mathbf{a}_r^2 \circ \cdots \circ \mathbf{a}_r^d$ is a rank-one tensor. R is the CP rank of the tensor $\widetilde{X}$. Correspondingly, the element $\widetilde{x}_{i_1, i_2, \ldots, i_d}$ in the tensor $\widetilde{X}$ is written as follows:

$$\begin{aligned} \widetilde{x}_{i_1, i_2, \ldots, i_d} &= \sum_{r=1}^{R} a_{i_1, r}^1 a_{i_2, r}^2 \ldots a_{i_d, r}^d, i_1 \\ &= 1, 2, \ldots, I_1, \ldots, i_d \\ &= 1, 2, \ldots, I_d, \end{aligned} \tag{2}$$

where $a_{i_k, r}^k$ is the element with an index of (i_k, r) in the factor matrix $A_k \in \mathbb{R}^{I_k \times R}$.

4. Dynamic Tensor-Based Bayesian Probabilistic Decomposition Approach

This section demonstrates a dynamic tensor-based Bayesian probabilistic decomposition approach for estimating the urban traffic states. As a common indicator of evaluating the traffic states, the average speed of the vehicle on a link

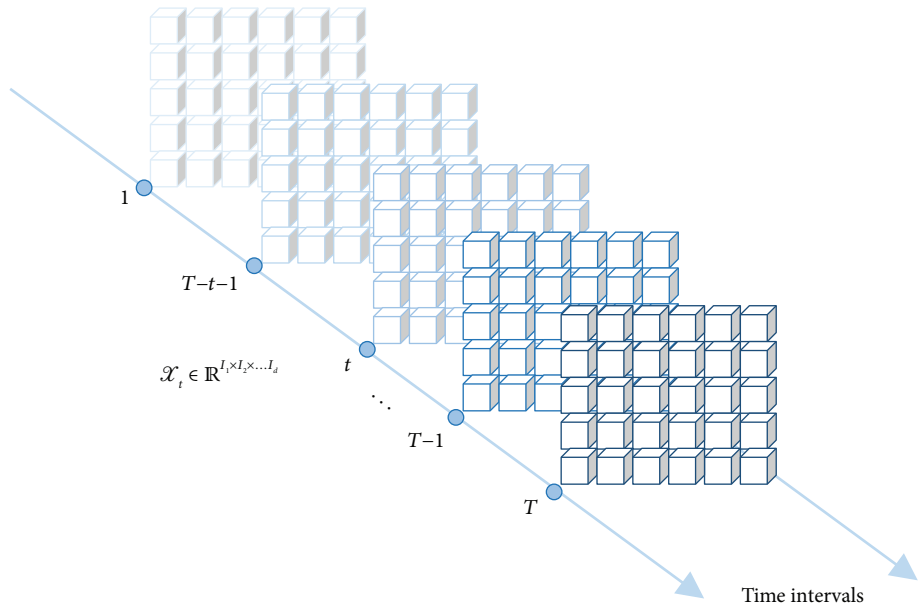

Figure 1: The schematic diagram of a d^{th}-order tensor stream.

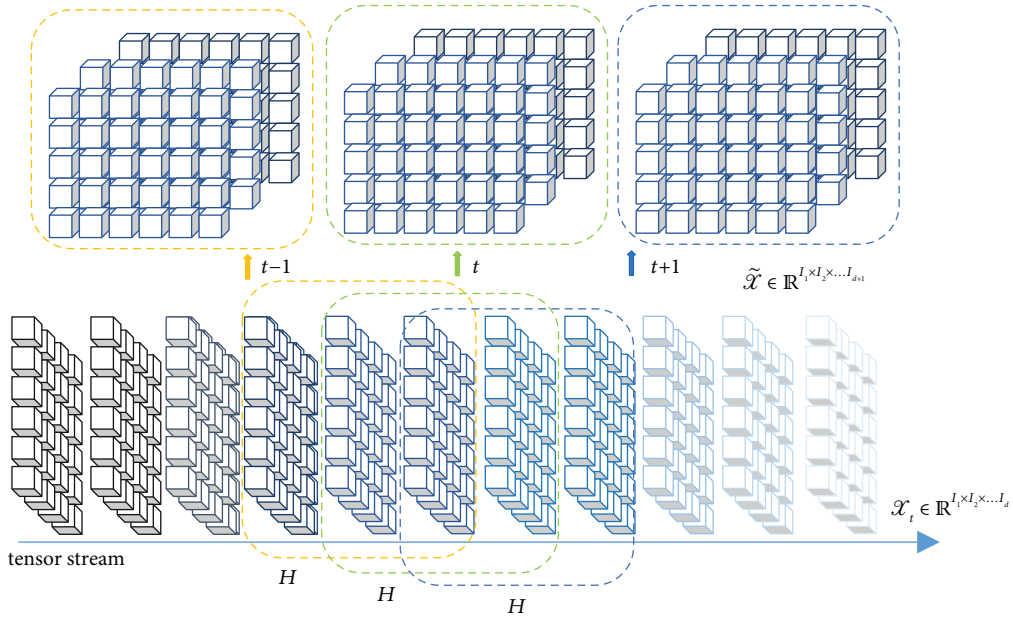

Figure 2: The divided tensor stream with a time horizon.

section during a time interval is utilized to represent the traffic state.

4.1. Problem Statement. Road network speed estimation is to predict the traffic speed in the future by using the speed data of the existing historical time.

Define $\overline{X} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_n}$ as the n_{th}-order tensor flow represented by road network speed data. $I_1, \ldots, I_n$ represent different dimensions of the tensor, such as the number of link sections, the number of days in different weeks, the number of time intervals, and other dimensions. The dynamic tensors $\widetilde{X} \in \mathbb{R}^{I_1 \times \ldots \times H \times \ldots \times I_n}$ can be extracted from the tensor stream with the sliding time window set as H. Note that $\widetilde{X} \in \mathbb{R}^{I_1 \times \ldots \times H \times \ldots \times I_n}$ contains the traffic speeds to be estimated with the prediction horizon set as P and $1 \le P \le H$.

If I_i represents the sliding time window of the dynamic tensor, the value of the I_i can be represented as H, which includes the traffic speeds of the time intervals that need to be estimated. Assume P defines the estimation horizon and $1 \le P \le H$; therefore, the state estimation can be written as follows:

$$\overline{\mathrm{y}}^{I_n \times P} = f\left(\widetilde{X}^{I_1 \times \ldots \times H \times \ldots \times I_n}\right), \tag{3}$$

where $\overline{\mathrm{y}}$ represents the traffic flow data to be estimated. $f(.)$ denotes the estimation model.

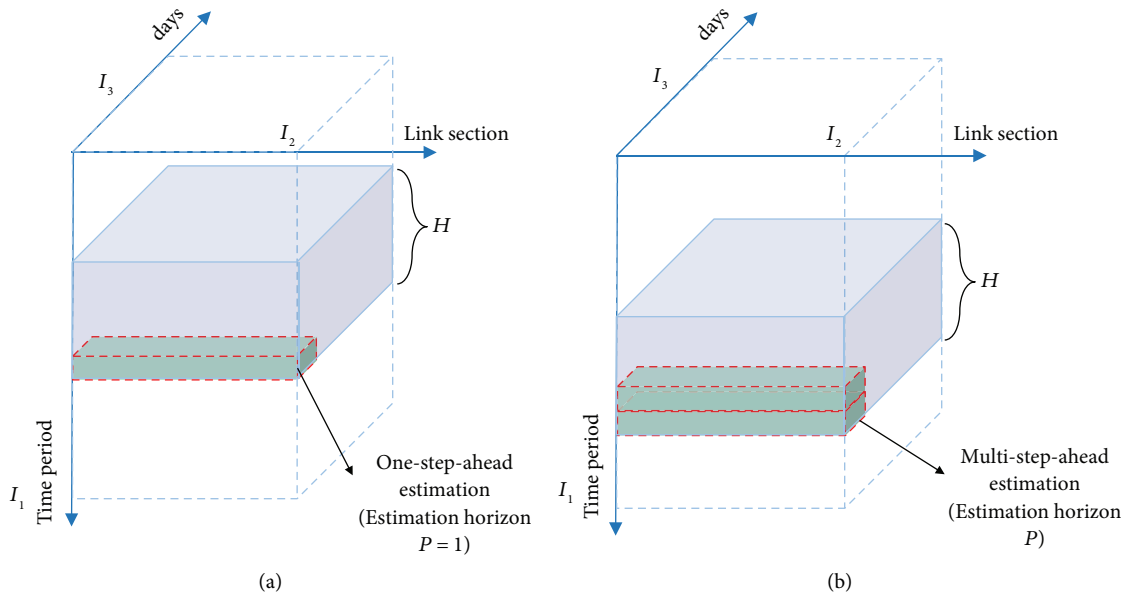

FIGURE 3: Example of the dynamic tensor with traffic speeds to be estimated. (a) One-step ahead estimation; (b) multistep ahead estimation.

Figure 3 shows the example of the dynamic tensor extracted from the tensor stream with three dimensions, where I_1 represents the periods, I_2 represents the number of the research link sections, and I_3 represents the number of days in different weeks. In addition, Figure 3 indicates two cases including the one-step-ahead prediction and multi-step-ahead estimation, respectively. Note that the iterated strategy indicated by Zhan et al. [55], which takes the estimated value as the input, was utilized to realize multistep ahead estimation.

4.2. Tensor-Based Bayesian Probabilistic CP Decomposition. This section introduces the Bayesian probabilistic CP decomposition utilized to produce the predicted value in a dynamic tensor $\widetilde{X}$. Assume $\widetilde{X}$ is a 3rd-order tensor and according to the definition of the CP decomposition, we have the following:

$$\widetilde{\mathbf{X}} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r, \tag{4}$$

where $\mathbf{a}_r$, $\mathbf{b}_r$, and $\mathbf{c}_r$ represent the r-th row of the factor matrix $\mathbf{A} \in \mathbb{R}^{I \times R}$, $\mathbf{B} \in \mathbb{R}^{J \times R}$, and $\mathbf{C} \in \mathbb{R}^{K \times R}$.

Suppose that the noise term of the measured entry $p \in P_o$ in the approximation obeys independent Gaussian distribution as follows:

$$x_{\mathbf{p}} \sim \mathbf{N}\left(\widehat{x}_{\mathbf{p}}, \sigma^{-1}\right), \tag{5}$$

where $\mathbf{N}(.)$ stands for the Gaussian distribution and σ is the precision item which is equal to the inverse of the covariance. Note that σ is a universal parameter for all elements.

To construct model on the tensor data adequately, flexible prior distributions are placed over both the precision item σ and the factor matrices. To be specific, the prior distributions over the row vectors in all factor matrices are assumed to obey multivariate Gaussians.

$$\begin{aligned} \mathbf{a}_i &\sim \mathbf{N}\left(\mu_a, (\Lambda_a)^{-1}\right), i = 1, 2, \ldots, I, \\ \mathbf{b}_j &\sim \mathbf{N}\left(\mu_b, (\Lambda_b)^{-1}\right), j = 1, 2, \ldots, J, \\ \mathbf{c}_k &\sim \mathbf{N}\left(\mu_c, (\Lambda_c)^{-1}\right), k = 1, 2, \ldots, K. \end{aligned} \tag{6}$$

The key concept of the fully Bayesian scheme is to treat the hyperparameters including σ, $\theta_a \equiv \{\mu_a, \Lambda_a\}$, $\theta_b \equiv \{\mu_b, \Lambda_b\}$, and $\theta_c \equiv \{\mu_c, \Lambda_c\}$ as random variables. According to the related work done by Salakhutdinov and Mnih [56], Tang et al. [54], and Chen et al. [52], conjugate Gaussian–Wishart priors can be set on the hyperparameters $\mu_a, \mu_b, \mu_c \in \mathbb{R}^{R \times 1}$ and $\overset{a}{\Lambda}, \overset{b}{\Lambda}, \overset{c}{\Lambda} \in \mathbb{R}^{R \times R}$ as follows:

$$\begin{aligned} p(\theta_a) &\sim \mathbf{N}\left(\mu_a, \left(\overset{0}{\beta}\Lambda_a\right)^{-1}\right) \times \mathbf{W}(\Lambda_a | W_0, \nu_0), \\ p(\theta_b) &\sim \mathbf{N}\left(\mu_b, \left(\overset{0}{\beta}\Lambda_b\right)^{-1}\right) \times \mathbf{W}(\Lambda_b | W_0, \nu_0), \\ p(\theta_c) &\sim \mathbf{N}\left(\mu_c, (\beta_0\Lambda_c)^{-1}\right) \times \mathbf{W}(\Lambda_c | W_0, \nu_0), \end{aligned} \tag{7}$$

where $\mathbf{G}(.)$ is a Gamma distribution. $\mathbf{W}(.)$ represents a Wishart distribution with ν_0 degrees of freedom and a $R \times R$ scale matrix W_0. $\mathbf{W}(.)$ can be written as follows:

$$\mathbf{W}\left(\Lambda^k | W_0, \nu_0\right) = \frac{1}{C}\left|\Lambda^k\right|^{(\nu_0 - R - 1)/2} \exp\left\{-\frac{1}{2} tr\left(W_0^{-1}\Lambda^k\right)\right\}, \tag{8}$$

where $tr(.)$ defines the trace function of a square matrix, which is the sum of all elements on its main diagonal.

Next, the precision item σ in (5) needs to be estimated. Under the assumption of Gaussian distribution, the parameters reflect the noise level. Under the scenario of urban traffic state estimation, the data is constantly updated with time, so the precision is unknown, and it is unable to be

completely captured by the reciprocal of the variance of all historical data. To improve the robustness of the model, the conjugate gamma distribution prior $\mathbf{G}(.)$ is set for the precision item and it can be expressed as follows:

$$\sigma \sim \mathbf{G}\left(\widetilde{W}_0, \widetilde{\nu}_0\right), \tag{9}$$

where $\widetilde{W}_0$ and $\widetilde{\nu}_0$ represent the shape parameter and the rate parameter, respectively.

And we have the following:

$$p(\sigma) \sim G\left(\widetilde{W}_0, \widetilde{\nu}_0\right) = \frac{\widetilde{\nu}_0^{\widetilde{W}_0} \sigma^{\widetilde{W}_0 - 1} \exp\left(-\widetilde{\nu}_0 \sigma\right)}{\Gamma\left(\widetilde{W}_0\right)}. \tag{10}$$

In this way, the distribution of the traffic speeds to be estimated in $\widetilde{X}$ can be obtained by the Bayesian inference algorithm as follows:

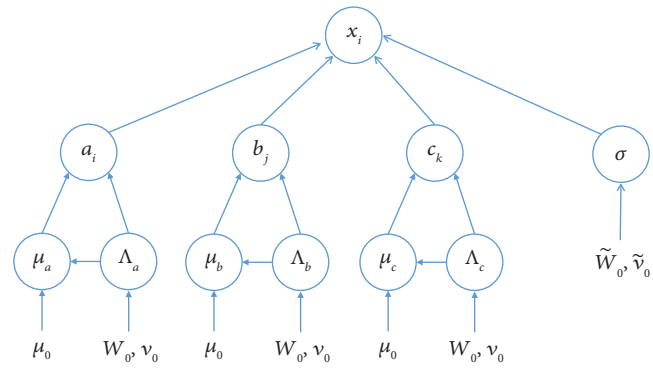

Figure 4: Graphical model for the DTBPD approach with 3rd-order tensor.

$$p\left(\widehat{x}_p|\widetilde{X}\right) = \int p\left(\widehat{x}_p|\mathbf{a}_i, \mathbf{b}_j, \mathbf{c}_k, \sigma\right) p\left(A, B, C, \sigma, \theta_a, \theta_b, \theta_c|\widetilde{X}\right) d\{A, B, C, \sigma, \theta_a, \theta_b, \theta_c\}, p \in P_e \tag{11}$$

Figure 4 indicates the graphical model which describes the generation procedure of the Bayesian probabilistic CP decomposition. Note that μ_0, $\widetilde{\nu}_0$, $\widetilde{W}_0$, ν_0, and W_0 are hyperparameters that should be given in advance.

4.3. Markov Chain Monte Carlo Model. Considering the distribution of the speeds in the near future by (11) is a multidimensional integral, numerical methods can be applied to solve this issue. By treating $p(\widehat{x}_p|\widetilde{X})$ as the expectation of $p(\widehat{x}_p|A, B, C, \sigma)$ over the posterior distribution $p(A, B, C, \theta_a, \theta_b, \theta_c, \sigma|\widetilde{X})$, $p(\widehat{x}_p|\widetilde{X})$ can be estimated by averaging the samples from this posterior distribution. To sample from the complicated posterior distribution that is hard to sample directly, a novel Gibbs sampling algorithm is provided to perform inference on the graphical model. Gibbs sampling algorithm is a widely used Markov Chain Monte Carlo-based approach for Bayesian inference. According to the computation details of conditional distribution discussed in reference [52, 54, 57], an efficient Gibbs sampling algorithm is summarized in Algorithm 1.

By running the Gibbs sampling algorithm several times under a stationary state, the estimation traffic speeds can be written as follows:

$$p\left(\widehat{x}_p|\widetilde{X}\right) = \frac{1}{M}\sum_{t=1}^{M} \mathbf{N}\left(\sum_{r=1}^{R} a_{ir}(t) b_{jr}(t) c_{kr}(t), {\sigma_t}^{-1}\right), \tag{12}$$

where $\widetilde{X}$ represents measured dynamic tensor. $\widehat{x}_p$ is an entry to be predicted and M is the number of iteration. $a_{jr}(t)$ is the (i, r) entry in the factor matrix A during the $t-$ th iteration. $b_{jr}(t)$ is the (i, r) entry in the factor matrix B during the $t-$ th iteration. $c_{kr}(t)$ is the (i, r) entry in the factor matrix C during the $t-$ th iteration. σ_t^{-1} defines precision item during the $t-$ th iteration.

5. Experiment

5.1. Trajectory Data Description and Preprocessing. The data used in this paper mainly include trajectory data and urban network data. The online car-hailing trajectory data of Xi'an provided by the Didi open project (https://outreach.didichuxing.com/research/opendata/) is employed to extract the trajectory data to capture the traffic speeds of the urban network. The network data were extracted from the open street map (OSM) to provide the map information of Xi'an city.

Since the trajectory data set used has the advantages of the high sampling frequency, wide coverage, and high precision, it is suitable for representing the traffic state. The original online car-hailing trajectory data covers about 60 square kilometers range of Xi'an city from Nov 1 to Nov 30 in 2016. The attributes of original datasets include driver ID, order ID, timestamp, longitude, latitude, and description. The sampling interval of trajectory is 2~4 seconds and the trajectory point can be mapped to the network.

Figure 5 illustrates the preprocessing of the original trajectory data, which can obtain the initial experiment datasets. As shown in Table 1, the initial experiment datasets include six attributes including the Driver_ID, Order_ID, Time Label, Longitude, Latitude, and Velocity.

5.2. Experiment Data and Analysis. To ensure the effectiveness of the experiment, we processed road network speed data for four weeks from Nov 1 to Nov 28 in 2016 as the experimental data. The data of the first 23 days were taken as historical data, and the data of the last 5 days were taken as testing data. Figure 6 provides the research scope of the experiment. As shown in Figure 6, the research scope area includes 64 roads including types of expressways, arterial roads, and secondary roads. To ensure a sufficient number of vehicles at each time interval and avoid inaccuracy or

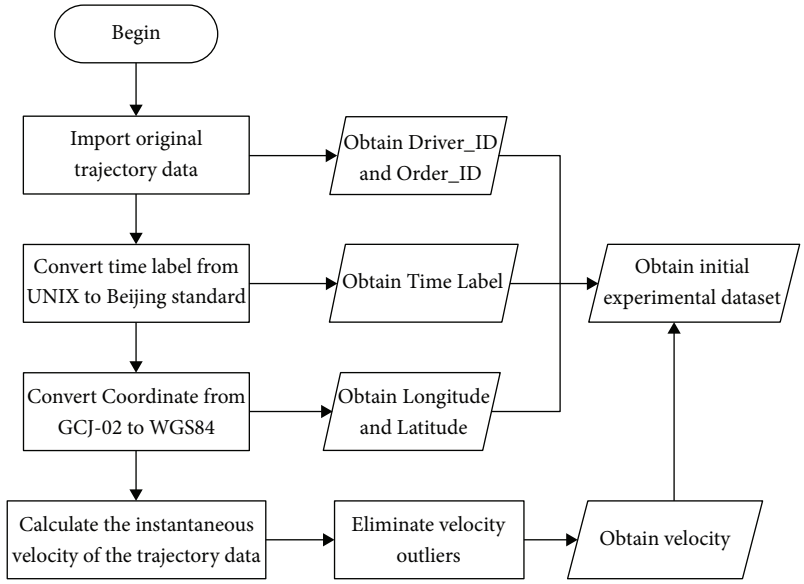

Figure 5: The preprocessing of the original trajectory data.

Initialize model parameter $\{A^{(1)}, B^{(1)}, C^{(1)}\}$
For $n = 1,2, \ldots, N$
Sample the hyper-parameter $\theta_a, \theta_b, \theta_c, \sigma$,respectively
$\sigma^{(n)}$: $p(\sigma^{(n)}|A^{(n)}, B^{(n)}, C^{(n)}, \bar{X})$
θ_a: $p(\theta_a|A^{(n)})$
θ_b: $p(\theta_b|B^{(n)})$
θ_c: $p(\theta_c|C^{(n)})$
For each $i = 1, 2, \ldots, I$, sample the time feature in parallel:
a_i^{n+1}: $p(a_i|B^{(n)}, C^{(n)}, \theta_a, \sigma^{(n)}, \bar{X})$
For each $j = 1, 2, \ldots, J$, sample the link section feature in parallel:
b_j^{n+1}: $p(b_j|A^{(n)}, C^{(n)}, \theta_b, \sigma^{(n)}, \bar{X})$
For each $k = 1, 2, \ldots, K$, sample the days in parallel:
c_k^{n+1}: $p(c_k|A^{(n)}, B^{(n)}, \theta_c, \sigma^{(n)}, \bar{X})$

Algorithm 1: Gibbs sampling for the DTBPD approach.

missing data caused by too long time interval, the travel speeds of the online car-hailing vehicle were average at each research road with a time interval of 5 min and each day is divided into 288 time intervals.

Figure 7 provides the average number of online car-hailing orders within the research area during different time intervals in a day. Considering the insufficient number of online car-hailing may lead to the inaccurate acquisition of traffic conditions, we select the periods from 8:00 to 22:00 when the number of orders is relatively large and sufficient taxis are running in the range of the research networks.

Figure 8 demonstrates the speed distribution of three different grades of roads in the study time range on a certain day. It can be seen from Figure 8 that in terms of speed distribution, the average speed of the expressway is the highest, and that of the secondary road is the lowest among the three types of roads, which is consistent with the functional design of the roads. In addition, the speeds of all roads decreased significantly during the period of 17:00 to 19:00 and the speeds of the expressways fluctuated most sharply among the three kinds of roads.

5.3. Spatiotemporal Analysis of Urban Traffic Data. Urban traffic flow data are complex and related to spatiotemporal dependency. Organizing the multidimensional traffic flow data into an appropriate tensor pattern is of importance to improve the efficiency and robustness of the traffic state estimation. In this section, the Pearson correlation coefficient is utilized to analyze the spatiotemporal characteristics of urban traffic. Since the residents usually travel regularly and urban traffic data often have certain periodic similarities, Figure 9 indicates the temporal dependency analysis of the urban roads from multiple perspectives. Based on the

Figure 6: The research scope area.

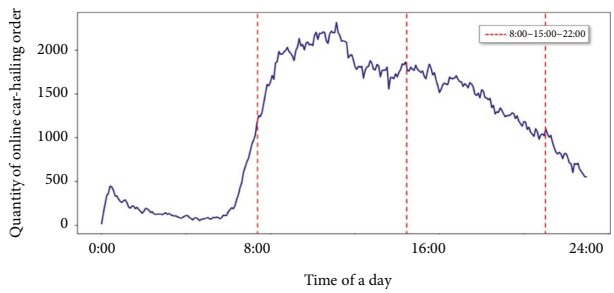

Figure 7: The average number of online car-hailing orders at time intervals in a day.

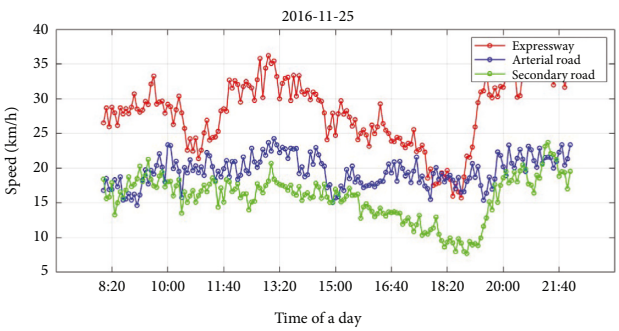

Figure 8: Comparison of the average speeds of the three kinds of roads in a day.

traffic speeds of a road link (ID = 63) during a day, Figure 9(a) illustrates that the traffic state of the same section varies greatly with the hours, indicating a shorter sliding time window may be suitable for dynamic tensor construction. As shown in Figure 9(b), the traffic state during days in a week shows a significant correlation with a correlation coefficient larger than 0.60. Meanwhile, Figures 9(c) and 9(d) demonstrate the high relevance of the traffic state among days in different weeks. As shown in Figures 9(c) and 9(d), these four days are all Thursday and the correlation coefficients among them are larger than 0.75. Hence, the dynamic tensor construction should consider more days with high relevance including the days in the same weeks and the days in different weeks.

Figure 10 indicates the spatial dependency analysis of the urban network by using the traffic speeds series data of all road links in the research scope during a day. It can be found that the road link on the network demonstrates spatial correlation in different degrees due to the connectivity of the urban road network. In general, the road links present high relevance with their upstream road section, downstream sections, or adjacent sections. Hence, this paper selects the highly correlated road links to construct dynamic tensors.

5.4. Baseline Model and Performance Indexes. To verify the effectiveness of the DTBPD model, several advanced deep learning-based models and tensor-based models are applied as the baseline models, including the LSTM [58], GRU [34], FDL [59], T-GCN [43], smooth PARAFAC tensor completion model (SPC) [60], HaLRTC [50], and BGCP model [52]. The parameters of the baseline models are indicated as follows:

LSTM: To solve the long-term dependency of RNN, the LSTM structure was established whose memory block contains forget gate, the input gate, and the output gate. To reduce overfitting issues, the dropout rate was set as 0.1.

GRU: As a variant of the LSTM with a simpler structure, the GRU model only combines the forget

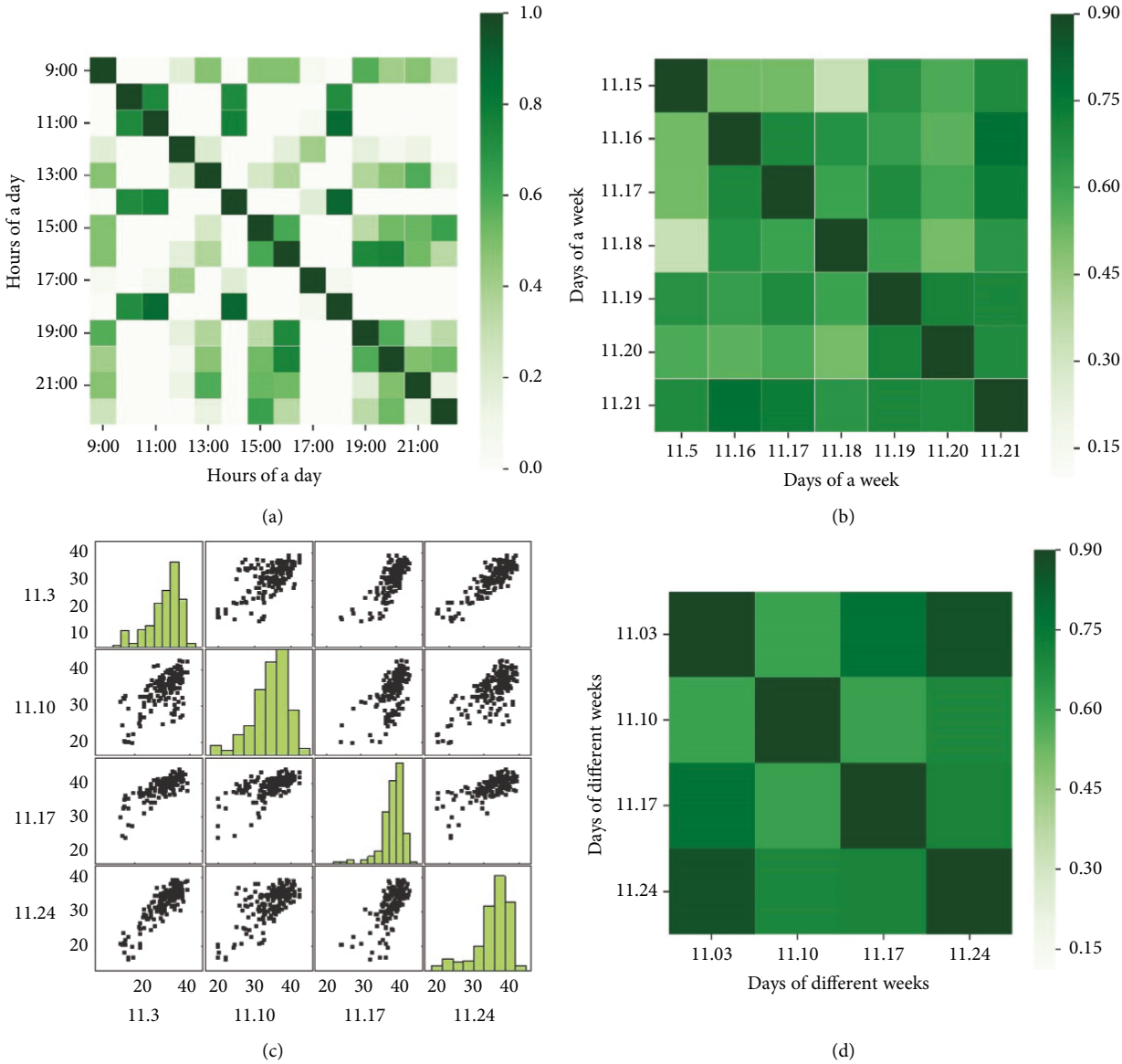

Figure 9: Temporal dependency analysis of the urban roads. (a) Correlation coefficient matrix of traffic state of different hours in a day; (b) correlation coefficient matrix of traffic state of different days in weeks; (c) speed scatter diagram of the Thursdays in different weeks; (d) correlation coefficient matrix of the Thursdays in different weeks.

gate and the input gate of the LSTM into a single update gate. Besides, there is a reset gate in a hidden unit of the GRU.

FDL: It is a combination model proposed by Gu et al. [59] and it fuses the LSTM and GRU layer, which can fully mine the spatiotemporal features with the selected input.

T-GCN: T-GCN [43] integrates the graph convolutional network and gated recurrent units to model spatial dependence and temporal dependence, respectively. There is one GRU layer in the hidden layers of the T-GCN model.

SPC: SPC is a smooth PARAFAC tensor-based completion model by combining the smooth PARAFAC decomposition and the efficient selection of models to minimize the tensor rank. Quadratic variation is employed as a strategy with parameter ρ chosen as [0.3 0.3 0.3].

HaLRTC: It is a high accuracy low-rank tensor completion approach employed by Ran et al. [50]. The parameter α is set as [1/3, 1/3, 1/3] and the parameter ρ is set as 0.01.

BGCP: It is a Bayesian decomposition-based tensor proposed by Chen et al. [52], which extends the Bayesian matrix factorization [56] to a higher-order case to learn the underlying statistical patterns in spatiotemporal traffic data.

In addition, the input time step of all the DL-based models is set as 5 and the input variable includes 64 link sections. For the LSTM, GRU, FDL, and T-GCN, the hidden

units in hidden layers are chosen as 200 and the epochs are all set as 200 with the batch size set as 64. To test the models fairly, the optimizer of three DL-based models is chosen as the Adam. For the tensor-based model, the input tensor contains the historical data of all road sections during the days in the week and the same days in the previous 4 weeks, indicating the dynamic tensors' dimension of the days and road sections 11 and 64, respectively. The low rank of the DTBPD algorithm is set as 30 and the MCMC-based sampling algorithm is run for 500 iterations. Note that the data of the first 23 days are utilized for training the DL-based model with the data of the last 5 days for testing these models.

To test the performance of the DTBPD approaches and compare it to the baseline models, the mean absolute percentage error (MAPE) and root mean square error (RMSE) were used to analyze the experimental results.

$$MAPE = \frac{1}{k}\sum_{n=1}^{k}\left|\frac{\widehat{\varphi}_k - \varphi_k}{\varphi_k}\right| \times 100\%,$$
$$RMSE = \sqrt{\frac{1}{k}\sum_{n=1}^{k}\left(\widehat{\varphi}_k - \varphi_k\right)^2}, \tag{13}$$

where $\widehat{\varphi}_k$ represents the estimated speed value and φ_k represents the ground-truth value. k is the number of testing speed data.

We utilize computer with Intel(R) Core(TM) i7-11800H CPU @ 4.6 GHz and 32 GB memory to conduct the experiment. Python 3.68 with TensorFlow 2.0 and Keras 2.0 is employed to implement the proposed model and baseline models.

6. Result and Discussion

6.1. Comparison of the DTBPD Approaches with Different Parameters. In this section, we test the DTBPD models by tuning up the low rank and the sliding time window, which are the two critical parameters of the DTBPD approach. We select the low rank from [5, 15, 30, 50, 100]. The sliding time windows are chosen from the set [6, 12, 24, 36, 48], which corresponds to the time of [0.5 h, 1 h, 2 h, 3 h, 4 h]. It can be learned from Figure 9, the MAPE and RMSE of the DTBPD model decrease with the increase of the low rank. As shown in Figure 11, the RMSE does not improve obviously with a large low rank. Hence, we set the low rank of the DTBPD as 30. In addition, the performance of the proposed model improves firstly and then drops with the increase of the sliding time window. This may be because the temporal feature of the traffic speed may drop if the distance between the historical data and the speed to be forecasted becomes larger, which may introduce more invalid temporal information and affect the performance of the DTBPD model. Since the proposed model can mine the characteristics of the traffic flow effectively, the sliding time window can be set as 12 which is most efficient for estimating traffic state in the future.

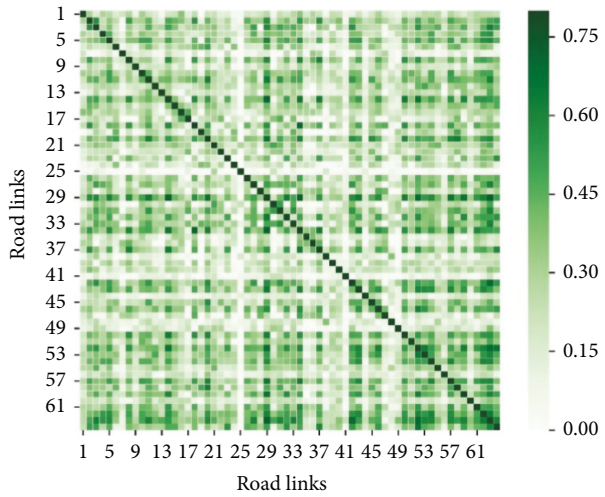

Figure 10: Spatial dependency analysis of the urban roads.

Figure 12 demonstrates the average MAPE and RMSE values of the DTBPD model at each time interval under different sliding time windows. It can be observed from Figure 10 that estimation errors of the starting period of the day are larger than those of other periods, which may be due to the composition of tensors considering the discontinuous traffic flow from the previous day. Figure 12 reveals that change trends of MAPE and RMSE are similar under different sliding time windows, indicating that the model is stable for different data ranges.

6.2. Performance Comparison of the Different Models. This section compares the overall performance of the DTBPD model with the baselines by taking the five-day estimation results as a whole. Tables 2 and 3 provide the MAPEs and RMSEs of the model at different types of roads. To be specific, the estimation error of all models at the expressways is little larger than those of the arterial road and secondary road. This may be since the values of the traffic speeds at expressways are relatively larger than those of the arterial roads and secondary roads, and the fluctuation of the traffic speeds at expressways are more violent. From the performance of the different models, it can be learned from Tables 2 and 3, the DTBPD outperforms the DL-based models and the tensor-based models, with improvements of 7.49% and 9.25% on MAPE and RMSE compared with those of the T-GCN model, which is the best DL-based predictor among the models. Since the T-GCN model combines the GCN and the GRU model, it shows better performance in learning the spatiotemporal characteristics than RNN-based models including the LSTM, GRU, and FDL. Meanwhile, the SPC and HaLRTC model also provides accurate and stable performance at the expressway and arterial road compared with the LSTM and GRU. Besides, the DTBPD model comes best among these models, which outperforms the BGCP model with improvements of 6.68% and 7.34% on MAPE and RMSE, respectively. Since the proposed model utilizes the dynamic input structure and Bayesian probabilistic CP decomposition, the traffic feature with high dependency can

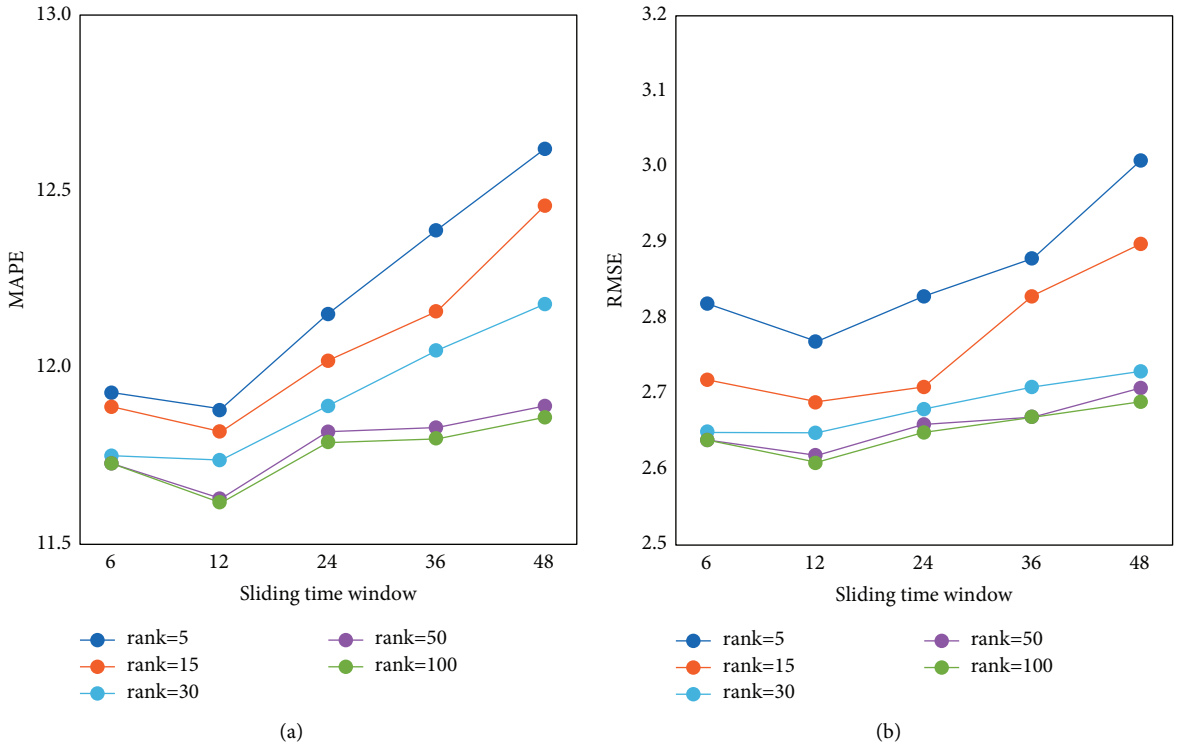

Figure 11: Comparison of the DTBPD model with different parameters. (a) MAPE; (b) RMSE.

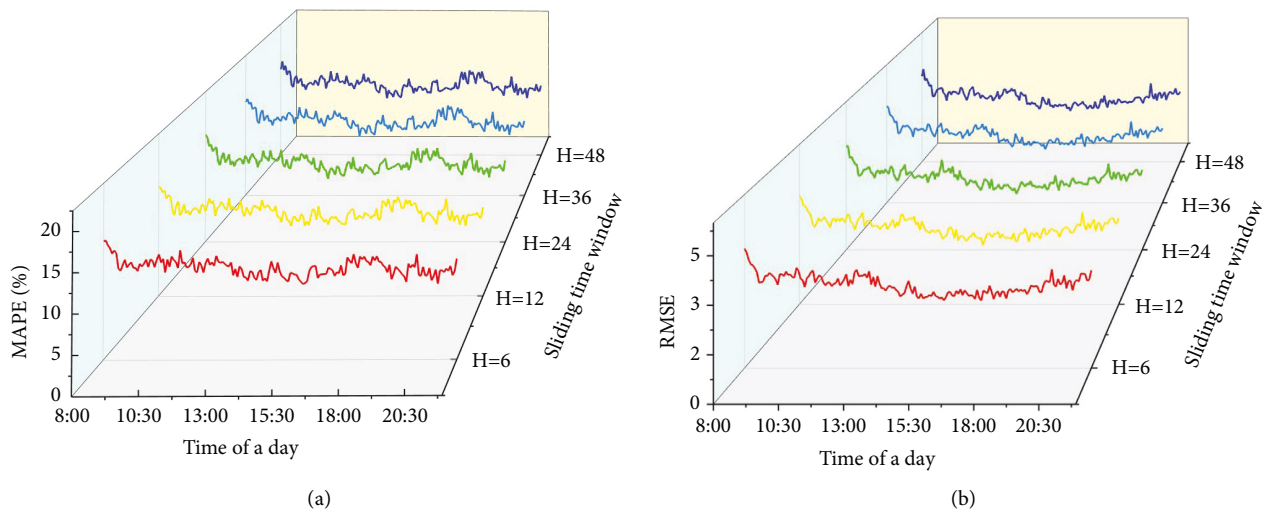

Figure 12: Average estimation speed of different time intervals with different sliding time windows. (a) MAPE; (b) RMSE.

Table 1: Examples of initial experimental datasets.

Driver_ID	Order_ID	Time Label	Longitude	Latitude	Velocity
0f8...322	5d6...891	2016/11/1 18:45:56	108.9425396	34.2667061	25.28
0f8...322	5d6...891	2016/11/1 18:45:59	108.9425496	34.2669360	30.62
0f8...322	5d6...891	2016/11/1 18:46:02	108.9425496	34.2670260	11.97
0f8...322	5d6...891	2016/11/1 18:46:05	108.9425495	34.2671659	18.63
0f8...322	5d6...891	2016/11/1 18:46:14	108.9425495	34.2674858	15.97
0f8...322	5d6...891	2016/11/1 18:46:17	108.9425495	34.2673958	11.97

Table 2: The overall performance of different models in terms of MAPE (%).

Models	Different types of roads			
	Expressway	Artertial road	Secondary road	Total
LSTM	13.87	13.72	13.32	13.64
GRU	13.54	13.03	12.88	13.16
FDL	13.04	12.98	12.58	12.88
T-GCN	12.82	12.71	12.44	12.69
SPC	13.85	13.67	13.45	13.65
HaLRTC	13.93	13.81	13.49	13.78
BGCP	12.77	12.78	12.03	12.58
DTBPD	**11.86**	**11.83**	**11.51**	**11.74**

Table 3: The overall performance of different models in terms of RMSE.

Models	Different types of roads			
	Expressway	Artertial road	Secondary road	Total
LSTM	3.42	3.29	3.13	3.32
GRU	3.26	3.18	3.12	3.19
FDL	3.19	3.13	3.03	3.12
T-GCN	3.03	2.99	2.82	2.92
SPC	3.36	3.25	3.02	3.20
HaLRTC	3.46	3.23	3.11	3.26
BGCP	2.92	2.94	2.74	2.86
DTBPD	**2.73**	**2.76**	**2.40**	**2.65**

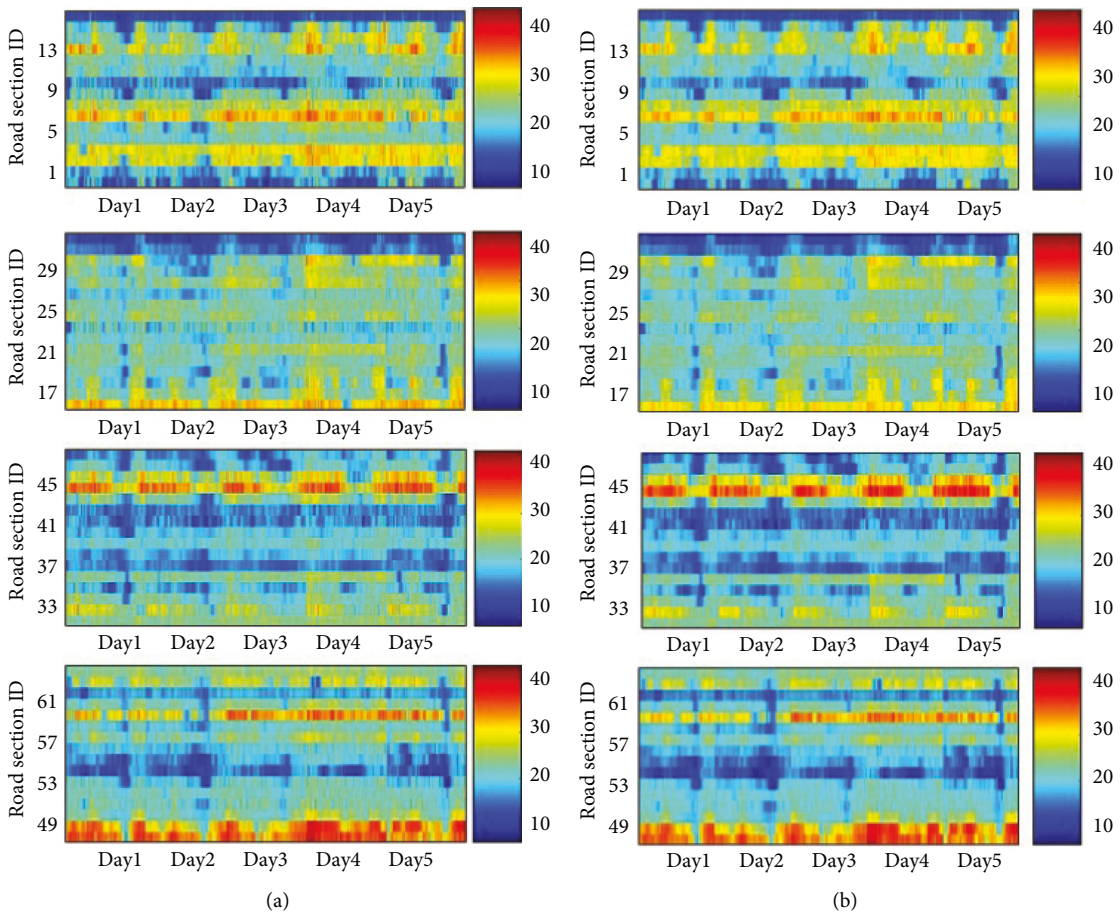

Figure 13: Comparison between the ground-truth and estimated data during the 5 days. (a) Ground-truth data; (b) estimated data.

Table 4: Comparison of time consumption between the DTBPD model and DL-based models.

Model	Training time consumption or calibration time consumption (s)
LSTM	380.35
GRU	310.40
FDL	774.08
T-GCN	823.31
DTBPD	2.49

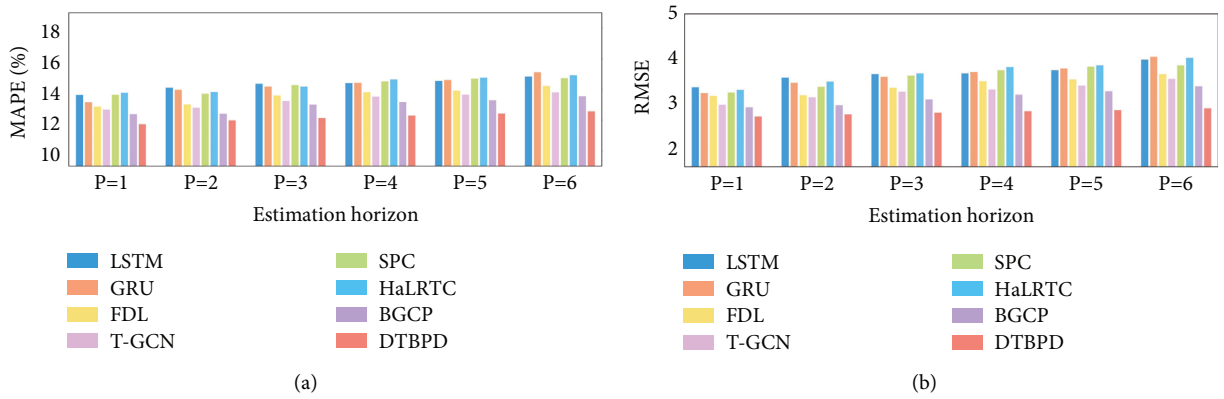

Figure 14: Comparison of different models with longer estimation horizons. (a) MAPE; (b) RMSE.

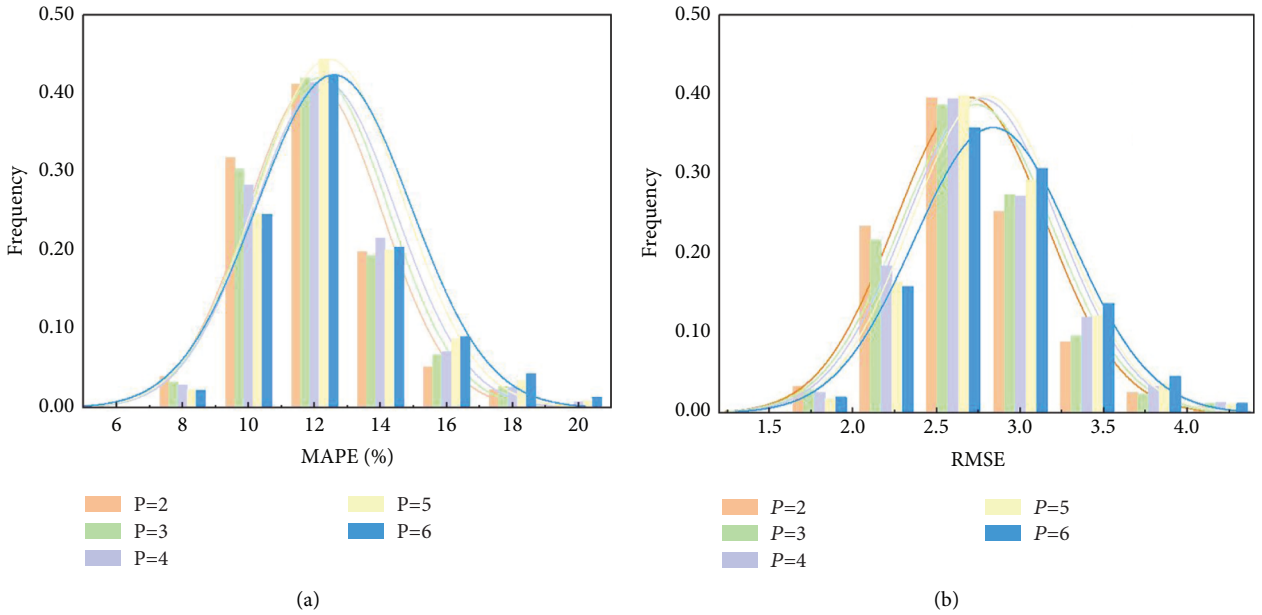

Figure 15: The error distribution of the DTBPD model with the increase of the estimation horizon. (a) MAPE; (b) RMSE.

be organized and mined in the tensor structure constantly. With the efficient decomposition approach, the calibrated factor matrices are capable of estimating the missing values in the dynamic tensor accurately.

Figure 13 gives the performance comparison between the ground-truth data and estimated data under the single-step-ahead estimation horizons. From the heat map, it is observed that the DTBPD model has a very good fitness

under normal and abnormal scenarios at different link sections. Note that even during the peak period when the traffic speed fluctuates significantly, the DTBPD can still estimate the traffic speed accurately. Meanwhile, even in different types of link sections where the traffic state presents differentiated features, the DTBPD model is capable of presenting similar estimation performance by using the dynamic tenor to consider the relative temporal characteristics and using the Gibbs sampling approach to approximate the factor matrix of the tensor.

Table 4 demonstrates the time consumption of the DTBPD model and the DL-based models. Since the DTBPD model needs to estimate the parameters of the model at each step, Table 4 provides its average time consumption for a single step. For the DL-based model, the update parameter in the model structure needs to train it with historical data, which spends much more time than the DTBPD model. Hence, compared with the DL-based models, the simple structure of the proposed model only needs a few seconds to calibrate the model, which is flexible and time-saving. In addition, the proposed model can adjust the composition of tensor data dynamically by using the sliding time window, which may protect the model from abnormal data and make the model robust.

6.3. Comparison of Different Models with Longer Estimation Horizons. Figure 14 provides the comparison of the different models with the extension of the estimation horizon from 5 mins to 30 mins. As shown in Figure 14, the MAPE and RMSE of different models rise slightly with the increase of the estimation horizon. Note that the DTBPD model still indicates obvious advantages over the LSTM, GRU, FDL, T-GCN, SPC, HaLRTC, and BGCP models when the estimation horizon $P = 6$, which means that this tensor-based model can efficiently infer the evolution trend of the traffic state.

Figure 15 provides the distributions of the MAPE and the RMSE estimated by the BTBPD approach with the 5-day dataset. As indicated in Figure 15, the errors of the proposed model are generally normal distribution. When the number of the estimation horizon increases from 2 to 6, the fitting curve moves to the right gradually and the prediction error gradually increases. This may be because the more step to estimate, the more estimated values in the historical data may enlarge the estimated error [55].

7. Conclusion

Massive online car-hailing trajectory data have become a popular and significant source for analyzing the urban traffic status, which is fundamental for reducing traffic congestion and constructing smart cities. This paper proposes a dynamic tensor-based Bayesian probabilistic decomposition approach for urban traffic state estimation. Firstly, the urban traffic speed data are formed into the dynamic tensor mode to fully mine the spatiotemporal characteristics of the traffic state. Secondly, the Bayesian probabilistic decomposition approach is introduced to decompose the tensor with speed to be estimated into a product of several vectors. Thirdly, the Gibbs sampling algorithm is proposed to calibrate the parameter of the proposed DBTPD models. Finally, the urban traffic speed data were extracted from the online car-hailing trajectory data of Didi Company to examine the accuracy and robustness of the DBTPD model.

Some findings can be summarized as follows. (1) The DBTPD model can well capture the spatiotemporal characteristics of the urban traffic state with a sliding sampling structure and Bayesian probabilistic decomposition approach. (2) The proposed outperforms the benchmark model including the LSTM, GRU, FDL, T-GCN, SPC, HaLRTC, and BGCP model in terms of the one-step-ahead estimation and multistep-ahead estimation. (3) The calibration time consumption of the DBTPD model is much less than those of the DL-based model, which indicates the proposed model is practical.

Some further research areas should be investigated. On the one hand, multisource data such as geographic information data or pedestrian flow data can be introduced to improve the accuracy of data estimation. On the other hand, the Bayesian probabilistic decomposition approach can be applied to multidimensional tensors, which consider more spatiotemporal characteristics for enhancing the robustness of the DBTPD model.

Data Availability

The trajectory data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

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

Authors' Contributions

Wenqi Lu was responsible for conceptualization, investigation, methodology, software, and writing—original draft preparation and review and editing. Ziwei Yi was responsible for visualization, methodology, and writing—original draft preparation. Dongyu Luo performed formal analysis and data curation. Yikang Rui was responsible for conceptualization, validation, and project administration. Bin Ran was responsible for project administration and supervision. Jianqing Wu performed formal analysis. Tao Li reviewed and edited the manuscript.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant no. 41971342), the Key Research and Development Program of Shandong Province (Grant no. 2020CXGC010118), the Fundamental Research Funds for the Central Universities, China (Grant no. 3221002145D), and the Scientific Research Foundation of Graduate School of Southeast University, China (Grant no. YBPY2161).

References

[1] J. Lu, B. Li, H. Li, and A. Al-Barakani, "Expansion of city scale, traffic modes, traffic congestion, and air pollution," *Cities*, vol. 108, Article ID 102974, 2021.

[2] X. Qian, T. Lei, J. Xue, Z. Lei, and S. V. Ukkusuri, "Impact of transportation network companies on urban congestion: evidence from large-scale trajectory data," *Sustainable Cities and Society*, vol. 55, Article ID 102053, 2020.

[3] Z. Chen, X. Lin, Y. Yin, and M. Li, "Path controlling of automated vehicles for system optimum on transportation networks with heterogeneous traffic stream," *Transportation Research Part C: Emerging Technologies*, vol. 110, pp. 312–329, 2020.

[4] D. Tian, G. Wu, P. Hao, K. Boriboonsomsin, and M. J. Barth, "Connected vehicle-based lane selection assistance application," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 7, pp. 2630–2643, 2019.

[5] L. Chen, A. H. Valadkhani, and M. Ramezani, "Decentralised cooperative cruising of autonomous ride-sourcing fleets," *Transportation Research Part C: Emerging Technologies*, vol. 131, Article ID 103336, 2021.

[6] E. Jenelius and H. N. Koutsopoulos, "Travel time estimation for urban road networks using low frequency probe vehicle data," *Transportation Research Part B: Methodological*, vol. 53, pp. 64–81, 2013.

[7] E. Jenelius and H. N. Koutsopoulos, "Probe vehicle data sampled by time or space: consistent travel time allocation and estimation," *Transportation Research Part B: Methodological*, vol. 71, pp. 120–137, 2015.

[8] T. T. Tchrakian, B. Basu, and M. O'Mahony, "Real-time traffic flow forecasting using spectral analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 519–526, 2011.

[9] Z. Mingheng, Z. Yaobao, H. Ganglong, and C. Gang, "Accurate multisteps traffic flow prediction based on SVM," *Mathematical Problems in Engineering*, vol. 2013, pp. 1–8, Article ID 418303, 2013.

[10] Y. Zhang, M. Lu, and H. Li, "Urban traffic flow forecast based on FastGCRNN," *Journal of Advanced Transportation*, vol. 2020, pp. 1–9, Article ID 8859538, 2020.

[11] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-term traffic forecasting: where we are and where we're going," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 3–19, 2014.

[12] D. Sun, K. Zhang, and S. Shen, "Analyzing spatiotemporal traffic line source emissions based on massive didi online car-hailing service data," *Transportation Research Part D: Transport and Environment*, vol. 62, pp. 699–714, 2018.

[13] J. Chen, W. Li, H. Zhang et al., "GPS data in urban online ride-hailing: a simulation method to evaluate impact of user scale on emission performance of system," *Journal of Cleaner Production*, vol. 287, Article ID 125567, 2021.

[14] B. Zhang, S. Chen, Y. Ma, T. Li, and K. Tang, "Analysis on spatiotemporal urban mobility based on online car-hailing data," *Journal of Transport Geography*, vol. 82, Article ID 102568, 2020.

[15] H. Tan, G. Feng, J. Feng, W. Wang, Y.-J. Zhang, and F. Li, "A tensor-based method for missing traffic data completion," *Transportation Research Part C: Emerging Technologies*, vol. 28, pp. 15–27, 2013.

[16] Z. Ma, H. N. Koutsopoulos, L. Ferreira, and M. Mesbah, "Estimation of trip travel time distribution using a generalized Markov chain approach," *Transportation Research Part C: Emerging Technologies*, vol. 74, pp. 1–21, 2017.

[17] W. Lu, T. Zhou, L. Li, Y. Gu, Y. Rui, and B. Ran, "An improved tucker decomposition-based imputation method for recovering lane-level missing values in traffic data," *IET Intelligent Transport Systems*, vol. 16, no. 3, pp. 363–379, 2022.

[18] Y. Han and Z. He, "Simultaneous incomplete traffic data imputation and similarity pattern discovery with bayesian nonparametric tensor decomposition," *Journal of Advanced Transportation*, vol. 2020, pp. 1–10, Article ID 8810753, 2020.

[19] M. S. Ahmed and A. R. Cook, "Analysis of freeway traffic time-series data by using box-jenkins techniques," *Transportation Research Record*, vol. 722, pp. 1–9, 1979.

[20] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results," *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 664–672, 2003.

[21] A. Emami, M. Sarvi, and S. A. Bagloee, "Short-term traffic flow prediction based on faded memory Kalman Filter fusing data from connected vehicles and Bluetooth sensors," *Simulation Modelling Practice and Theory*, vol. 102, Article ID 102025, 2020.

[22] H. P. Lu, Z. Y. Sun, and W. C. Qu, "Big data-driven based real-time traffic flow state identification and prediction," *Discrete Dynamics in Nature and Society*, vol. 2015, Article ID 284906, 11 pages, 2015.

[23] Z. Song, F. Sun, R. Zhang, Y. Du, and C. Li, "Prediction of road network traffic state using the NARX neural network," *Journal of Advanced Transportation*, vol. 2021, pp. 1–17, Article ID 2564211, 2021.

[24] J. Rice and E. vanZwet, "A simple and effective method for predicting travel times on freeways," *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 3, pp. 200–207, 2004.

[25] M. Castro-Neto, Y.-S. Jeong, M.-K. Jeong, and L. D. Han, "Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6164–6173, 2009.

[26] B. Sun, W. Cheng, P. Goswami, and G. Bai, "Short-term traffic forecasting using self-adjusting k-nearest neighbours," *IET Intelligent Transport Systems*, vol. 12, no. 1, pp. 41–48, 2018.

[27] L. Li, L. Qin, X. Qu, J. Zhang, Y. Wang, and B. Ran, "Day-ahead traffic flow forecasting based on a deep belief network optimized by the multi-objective particle swarm algorithm," *Knowledge-Based Systems*, vol. 172, pp. 1–14, 2019.

[28] L. Li, B. Du, Y. Wang, L. Qin, and H. Tan, "Estimation of missing values in heterogeneous traffic data: application of multimodal deep learning model," *Knowledge-Based Systems*, vol. 194, Article ID 105592, 2020.

[29] Y. Rui, W. Lu, Z. Yi, R. Wu, and B. Ran, "A novel hybrid model for predicting traffic flow via improved ensemble learning combined with deep belief networks," *Mathematical Problems in Engineering*, vol. 2021, no. 2, pp. 1–16, Article ID 7328056, 2021.

[30] Y. Hou, Z. Deng, and H. Cui, "Short-term traffic flow prediction with weather conditions: based on deep learning algorithms and data fusion," *Complexity*, vol. 2021, pp. 1–14, Article ID 6662959, 2021.

[31] J. Liu, X. Wang, Y. Li, X. Kang, and L. Gao, "Method of evaluating and predicting traffic state of highway network based on deep learning," *Journal of Advanced Transportation*, vol. 2021, pp. 1–9, Article ID 8878494, 2021.

[32] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *IEEE*

Transactions on Intelligent Transportation Systems, vol. 16, 2015.

[33] X. Luo, D. Li, nt(ancestor::ref)"=2?^!"count(../name)"=1?^!"count(following-sibling::name)" and >Y. Yang, and S. Zhang, "Spatiotemporal traffic flow prediction with KNN and LSTM," *Journal of Advanced Transportation*, vol. 2019, pp. 1–10, Article ID 4145353, 2019.

[34] F. Rui, Z. Zuo, and L. Li, "Using LSTM and GRU neural network methods for traffic flow prediction," in *Proceedings of the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pp. 324–328, Wuhan, China, November 2016.

[35] H. Yan, X. Ma, and Z. Pu, "Learning dynamic and hierarchical traffic spatiotemporal features with transformer," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–14, 2021.

[36] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang, "Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction," *Sensors*, vol. 17, no. 4, 2017.

[37] L. Zhao, D. Wei, Y. Dongmei, C. Gan, and G. Jianhua, "Short-term traffic flow forecast based on combination of K nearest neighbor algorithm and support vector regression," *Journal of Highway and Transportation Research and Development*, vol. 34, no. 5, pp. 122–128, 2017.

[38] Y. Liu, H. Zheng, X. Feng, and Z. Chen, "Short-term traffic flow prediction with Conv-LSTM," in *Proceedings of the 2017 9th International Conference on Wireless Communications and Signal Processing*, pp. 1–6, Nanjing, China, October 2017.

[39] Y. Gu, W. Lu, X. Xu, L. Qin, Z. Shao, and H. Zhang, "An improved bayesian combination model for short-term traffic prediction with deep learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 3, pp. 1332–1342, 2020.

[40] S. Guo, Y. Lin, S. Li, Z. Chen, and H. Wan, "Deep spatial-temporal 3D convolutional neural networks for traffic data forecasting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3913–3926, 2019.

[41] Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang, "A hybrid deep learning based traffic flow prediction method and its understanding," *Transportation Research Part C: Emerging Technologies*, vol. 90, no. 5, pp. 166–180, 2018.

[42] L. Li, X. Sheng, B. Du, Y. Wang, and B. Ran, "A deep fusion model based on restricted Boltzmann machines for traffic accident duration prediction," *Engineering Applications of Artificial Intelligence*, vol. 93, Article ID 103686, 2020.

[43] L. Zhao, Y. Song, C. Zhang et al., "T-GCN: a temporal graph convolutional network for traffic prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3848–3858, 2020.

[44] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, "Attention based spatial-temporal graph convolutional networks for traffic flow forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 922–929, 2019.

[45] W. Lu, Z. Yi, W. Liu, Y. Gu, Y. Rui, and B. Ran, "Efficient deep learning based method for multi-lane speed forecasting: a case study in Beijing," *IET Intelligent Transport Systems*, vol. 14, no. 14, pp. 2073–2082, 2020.

[46] W. Chen, J. An, R. Li, and G. Xie, "Tensor-train fuzzy deep computation model for citywide traffic flow prediction," *IEEE Access*, vol. 7, Article ID 120581, 2019.

[47] M. Bhanu, J. Mendes-Moreira, and J. Chandra, "Embedding traffic network characteristics using tensor for improved traffic prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3359–3371, 2021.

[48] J. Liao, J. Tang, W. Zeng, and X. Zhao, "Efficient and accurate traffic flow prediction via incremental tensor completion," *IEEE Access*, vol. 6, Article ID 36897, 2018.

[49] B. Ran, L. Song, J. Zhang, Y. Cheng, and H. Tan, "Using tensor completion method to achieving better coverage of traffic state estimation from sparse floating car data," *PLoS One*, vol. 11, no. 7, Article ID e0157420, 2016.

[50] B. Ran, H. Tan, Y. Wu, and P. J. Jin, "Tensor based missing traffic data completion with spatial-temporal correlation," *Physica A: Statistical Mechanics and Its Applications*, vol. 446, pp. 54–63, 2016.

[51] E. Acar, D. M. Dunlavy, T. G. Kolda, and M. Mørup, "Scalable tensor factorizations for incomplete data," *Chemometrics and Intelligent Laboratory Systems*, vol. 106, no. 1, pp. 41–56, 2011.

[52] X. Chen, Z. He, and L. Sun, "A Bayesian tensor decomposition approach for spatiotemporal traffic data imputation," *Transportation Research Part C: Emerging Technologies*, vol. 98, pp. 73–84, 2019.

[53] K. Tang, C. Tan, Y. Cao, J. Yao, and J. Sun, "A tensor decomposition method for cycle-based traffic volume estimation using sampled vehicle trajectories," *Transportation Research Part C: Emerging Technologies*, vol. 118, Article ID 102739, 2020.

[54] K. Tang, S. Chen, Z. Liu, and A. J. Khattak, "A tensor-based Bayesian probabilistic model for citywide personalized travel time estimation," *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 260–280, 2018.

[55] X. Zhan, S. Zhang, W. Y. Szeto, and X. Chen, "Multi-step-ahead traffic speed forecasting using multi-output gradient boosting regression tree," *Journal of Intelligent Transportation Systems*, vol. 24, no. 2, pp. 125–141, 2020.

[56] R. Salakhutdinov and A. Mnih, "Bayesian probabilistic matrix factorization using Markov chain Monte Carlo," in *Proceedings of the 25th International Conference on Machine Learning*, pp. 880–887, Association for Computing Machinery, NY, USA, July 2008.

[57] L. Xiong, X. Chen, T.-K. Huang, J. Schneider, and J. G. Carbonell, "Temporal collaborative filtering with bayesian probabilistic tensor factorization," in *Proceedings of the 2010 SIAM international conference on data mining*, pp. 211–222, Milan, Italy, December 2010.

[58] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, 2015.

[59] Y. Gu, W. Lu, L. Qin, M. Li, and Z. Shao, "Short-term prediction of lane-level traffic speeds: a fusion deep learning model," *Transportation Research Part C: Emerging Technologies*, vol. 106, pp. 1–16, 2019.

[60] T. Yokota, Q. Zhao, and A. Cichocki, "Smooth PARAFAC decomposition for tensor completion," *IEEE Transactions on Signal Processing*, vol. 64, no. 20, pp. 5423–5436, 2016.