Short-Term Traffic Prediction Based on Dynamic Tensor Completion

Huachun Tan, *Member, IEEE*, Yuankai Wu, Bin Shen, *Member, IEEE*, Peter J. Jin, and Bin Ran

***Abstract*—Short-term traffic prediction plays a critical role in many important applications of intelligent transportation systems such as traffic congestion control and smart routing, and numerous methods have been proposed to address this issue in the literature. However, most, if not all, of them suffer from the inability to fully use the rich information in traffic data. In this paper, we present a novel short-term traffic flow prediction approach based on dynamic tensor completion (DTC), in which the traffic data are represented as a dynamic tensor pattern, which is able capture more information of traffic flow than traditional methods, namely, temporal variabilities, spatial characteristics, and multimode periodicity. A DTC algorithm is designed to use the multimode information to forecast traffic flow with a low-rank constraint. The proposed method is evaluated on real-world data sets and compared with other state-of-the-art methods, and the efficacy of the proposed approach is validated on the experiments of traffic flow prediction, particularly when dealing with incomplete traffic data.**

***Index Terms*—Short-term traffic flow prediction, multi-mode information, dynamic tensor completion, missing data.**

I. Introduction

TRAFFIC congestion has become increasingly severe worldwide. Due to the limitation of ability to increase the infrastructure and to restrict traveling, as an alternative strategy, the intelligent transportation systems which aim to enhance the efficiency of exit transportation system [1], [2] such as traffic network planning, route planning, and driver assistance systems [3], [4], have become vital for alleviating traffic congestion, while the prediction of near-term traffic flow is an important part for such system. For example, the information of short-term traffic flow prediction can be provided to drivers in real-time to give them realistic estimation of travel data, expected delays and alternative routes to their destinations. It is believed that providing drivers with this information can help alleviate traffic congestion and enhance the performance of the entire driver–vehicle–road networks [5].

Due to the importance of short-term traffic prediction, a large number of techniques for short-term traffic flow prediction have been proposed in the last few decades [6], [7]. Summing up these prediction models, short-term traffic prediction methods are usually based on three types of information sources: temporal variabilities, spatial characteristics and multi-mode periodicity.

Most of the prediction methods were proposed to utilize the temporal variabilities to forecast traffic flow, and the traffic data are always represented as time series. A common example of such methods is the ARIMA model [8]–[10], which assumes that the future traffic volume varies linearly according to the prevailing traffic. To capture the non-linear variation of traffic flow data, Smith *et al.* [11] propose the nonparametric models for traffic flow prediction; Castro-Neto *et al.* [12] use support vector regression to utilize the non-linear variation for traffic prediction. And many neural network models have been proposed to capture the past-future relationship from the training data [13]–[15]. However, for traffic flow, only using temporal variation is not sufficient for accurately forecasting, especially when it is for traffic flow prediction of multiple locations. To use more information, numerous methods incorporate the spatial characteristics of traffic flow, including the multivariate time series method [16]–[18] and neural networks with inputs of spatial information [19]. Generally, such methods construct traffic flow into 2-way pattern, which reflects both the temporal variabilities and spatial characteristics of traffic flow. Compared with prediction models only incorporated with temporal information, the spatial-temporal approaches usually achieve better performance. By the way, the 2-way data pattern also achieves better performance than 1-way pattern that only using temporal information for missing traffic data imputation [20]–[22].

Another basic property of traffic flow data is the temporal periodic characteristics. For example, there is strong periodicity over a 24-hour period for weekdays. Some research studies have indicated that prediction performance can be notably improved by utilizing these periodic characteristics [23]. To address these periodic characteristics, many methods considering periodicity have been proposed. Williams and Hoel [24] proposed seasonal ARIMA to forecast traffic flow and it

Manuscript received March 20, 2015; revised October 31, 2015 and November 28, 2015; accepted December 2, 2015. Date of publication February 16, 2016; date of current version July 29, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 6127137651308115, by the National Basic Research Program of China (973 Program) under Grant 2012CB725405, by the Beijing Natural Science Foundation under Grant 4122067, and by the Basic Research Fund of the Beijing Institute of Technology under Grant 20120342026. The Associate Editor for this paper was L. Li.

H. Tan and Y. Wu are with the Department of Transportation Engineering, Beijing Institute of Technology, Beijing 100081, China (e-mail: tanhc@bit.edu.cn; Kaimaogege@gmail.com).

B. Shen was with the Department of Computer Science, Purdue University, West Lafayette, IN 47907 USA. He is now with Google Research, New York, NY 10011 USA (e-mail: bshen@purdue.edu).

P. J. Jin is with the Department of Civil and Environmental Engineering, Rutgers State University of New Jersey, Piscataway, NJ 08854 USA (e-mail: peter.j.jin@rutgers.edu).

B. Ran is with the Department of Civil and Environmental Engineering, University of Wisconsin–Madison, Madison, WI 53706 USA (e-mail: bran@wisc.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2015.2513411

outperforms the ARIMA models by taking the seasonal trend of traffic flow into account. Hong [25] developed a seasonal SVR model to forecast traffic flow. A comparison and integrated study on periodicity modeling of traffic flow time series can be found in [26]. And several hybrid methods have been proposed, the popular and prevailing strategy is the combination of a conventional prediction models and a method that can capture the periodicity of data [14], [27], [28].

While all the previously mentioned prediction approaches are powerful and useful methods for traffic flow prediction, they also have some drawbacks facing with complex traffic flow data pattern that have a multi-mode structure(time, space, week, day, *et al.*) [29]–[31]. Generally, the more the multi-mode information prediction method incorporated, the more accurate the prediction method is. For example, prediction model with both spatial mode and temporal mode information outperforms the prediction model purely exploiting temporal information, and the performance of prediction method can be improved by utilizing the week mode and day mode periodicity. However, most of the conventional prediction methods construct traffic data into time series (1-way) or matrix pattern (2-way) [32]. The one-way vector and two way pattern only mine limited mode characteristics because the mode number of traffic flow data is more than the mode dimension of matrix pattern (2-way). In order to mine more multi-mode information under the umbrella of one-way and two-way data structure pattern, many research studies propose hybrid models to forecast short-term traffic flow. Although some hybrid methods achieve promising performance, the combination rule of different methods is difficult to determine and the combination may add complexity to the prediction model. In view of these deficiencies, this paper focuses on developing a method that can make full use of the multi-mode information of traffic data to improve short-term traffic flow prediction.

In the last decades, one of the most popular method to characterize the data with multi-mode structure is to construct such data into tensor (multi-way array), which is the multi-way generalizations of vector and matrix [33]–[35]. It has been shown in numerous research areas, including face recognition [36], social networks [37], video processing [38] and neuroscience [39]. Constructing data with multi-mode structure into tensor pattern with advantages over 1-way and 2-way pattern in terms of extension of the multi-mode structure utilization capability. Particularly, recent studies show that the tensor-based methods outperform 1-way and 2-way based methods for missing traffic data imputation [29] and recovery of traffic data with outliers [40]. Recently, there are attempts to develop tensor decomposition algorithm for traffic speed prediction. In [41], the tensor decomposition is utilized to provide a low-dimensional representation of traffic state. The solution of traffic prediction in [41] is a combination of tensor decomposition and SVR. It should be noted that the application of tensor decomposition in [41] was focused on the low-dimensional representation of traffic state. And this method did not consider the problem of missing data, which is a major issue in traffic data sets.

Motivated by the success of tensor-based methods, which consider the multi-mode information of traffic data in the area of traffic data processing, we adopt the dynamic tensor pattern to model traffic flow data for short-term traffic data and propose a novel dynamic tensor completion method (DTC). DTC can utilize the temporal mode, spatial mode, day mode and week mode information of traffic flow data. Also, we use a tensor completion method, which is believed to be one of the best multi-way unknown data estimation method [42], [44]. Consequently, the proposed model accurately forecasts the near-term traffic flow with the multi-mode information. For evaluation, the proposed method is tested on real-world data sets and compared with other classic prediction models. The experimental results show that the proposed DTC achieves satisfactory performance. In summary, this paper first proposes dynamic tensor completion (DTC) for forecasting the data with multi-mode structure, then a dynamic tensor completion algorithm using the multi-mode low-rank properties is proposed to estimate the unknown data, and it is further demonstrated that the proposed method can still accurately forecast traffic data on a data sets with missing data.

The main contributions of this paper include the following

1) We proposed a dynamic tensor completion method dealing with dynamic tensor stream constructed from time series with multi-mode structure. The tensor completion problem is solved by fast matrix factorization. It contrasts to our previous tensor completion models such as reported in [29], [40] and [45], where, static tensor, not dynamic tensor, model was used.
2) The dynamic tensor completion method is applied to short-term traffic flow prediction. The prediction is purely fulfilled by tensor completion without a training process of the predictors. And it can effectively deal with missing data issues.
3) We designed a dynamic traffic flow tensor model according to the multi-mode correlations of traffic flow data. And we use a method to dynamically find the suitable low-n-rank of the dynamic tensor model.

The rest of this paper is organized as follows. The necessary theoretical background is introduced in Section II. Section III presents the dynamic tensor completion based prediction method. Experiments are reported in Section IV to verify the applicability and effectiveness of our dynamic tensor completion method with real traffic data obtained from PeMS [47] System. Section V gives some further discussion of the proposed method. Then, in Section VI, we sum up our method and discuss the future research direction.

II. Theoretical Background

A. Tensor Basics

Multiway arrays, also referred to as tensors, are higher-order generalizations of vectors and matrices. Higher-order arrays are represented as $\mathbf{X} \in R^{I_1 \times I_2 \times, \ldots, I_N}$, where the order of $\mathbf{X}$ is N. Each dimension of a multiway array is called a mode [48]. The mode-n unfolding (also called matricization or flattening) of a tensor $\mathbf{X} \in R^{I_1 \times I_2 \times, \ldots, I_N}$ is defined as unfolding $(\mathbf{X}, n) = X_{(n)}$, where the tensor element $(i_1, i_2, \ldots, i_n)$ is mapped to the matrix element (i_n, j), where

$$j = 1 + \sum_{\substack{k=1\\k\neq n}}^{N} (i_k - 1) J_k \quad \text{with} \quad J_k = \prod_{m=1.m\neq n}^{k-1} I_m. \tag{1}$$

Therefore, $X_{(n)} \in \mathrm{R}^{I_N \times J}$, where $J = \prod_{k=1.k\neq n}^{k-1} I_k$. The n-rank of a N-dimensional tensor $\mathbf{X} \in R^{I_1 \times I_2 \times, \ldots, I_N}$, denoted by $R_{(n)}$, is the rank of the mode-n unfolding matrix $X_{(n)}$.

The inner product of two same-size tensors $\mathbf{A}, \mathbf{B} \in R^{I_1 \times I_2 \times, \ldots, I_N}$ is defined as the sum of the products of their entries, i.e.,

$$(\mathbf{A}, \mathbf{B}) = \sum_{i_1} \sum_{i_2} \cdots\cdots \sum_{i_n} a_{i_1, \ldots, i_k, \ldots, i_n} b_{i_1, \ldots, i_k, \ldots, i_n}. \tag{2}$$

The corresponding Frobenius norm is $\|\mathbf{X}\|_F = \sqrt{(\mathbf{X}, \mathbf{X})}$. For any $1 \leq n \leq N$, the n-mode (matrix) product of a tensor $\mathbf{A} \in R^{I_1 \times I_2 \times, \ldots, I_N}$ with a matrix $M \in R^{J \times I_n}$ is denoted by $\mathbf{A} \times_n M$. In terms of flattened matrix, the n-mode product can be expressed as

$$\mathbf{Y} = \mathbf{A} \times_n M \iff Y_{(n)} = M A_{(n)}. \tag{3}$$

Let $\mathbf{A}^{\mathbf{W}}$ be the $I_1 \times I_2 \times, \ldots, I_N$ observed tensor that stores all the observed values, such that

$$\mathbf{A}^{\mathbf{W}} = \begin{cases} a_{i_1, \ldots, i_k, \ldots, i_n} & \text{if } (i_1, \ldots, i_k, \ldots, i_n) \in \mathbf{W} \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

In tensor completion, a subset of the entries of the tensor $\mathbf{A}$ is given, and under the low-n-rank assumption, the unknown entries are deduced by taking advantage of a global property of the data [29], [42], [49], [50]:

$$\min_{\mathbf{A} \in \mathbf{T}} \sum_{i_1}^{N} \text{rank}\left(A_{(i)}\right) \quad \text{s.t.} \quad \mathbf{A}^{\mathbf{W}} = \mathbf{T}^{\mathbf{W}}. \tag{5}$$

B. Dynamic Tensor Model

Dynamic tensor is a data model to represent a multi-mode stream that grows incrementally over time, see [51] and [52] for example. The tensor stream can be defined as:

Tensor Stream: A sequence of m_{th}-order tensors $\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_T}$, where each $\mathbf{X_t} \in R^{I_1 \times I_2 \times, \ldots, I_m} (1 \leq t \leq T)$, is called a $m + 1_{th}$ order tensor stream if $\mathbf{T}$ is the maximum index that increases with time.

As Fig. 1 shows, the tensor stream is considered to be incrementally growing over time. And $\mathbf{X_T}$ is the latest tensor in the stream.

In many cases, it is not necessary to use all of the tensors in the tensor stream. A tensor window can be used to localize the tensor stream into a smaller tensor sequence with size W at time t, where the tensor window $\mathbf{D}(t, w) = \{\mathbf{X_{(T-w+1)}}, \ldots, \mathbf{X_T}\}$ with each $\mathbf{X_t} \in R^{I_1 \times I_2 \times, \ldots, I_{m+1}}$. Fig. 2 gives a abridged general view of sliding tensor window.

For the dimensionality reduction and feature extraction of the tensor stream, the tensor stream decomposition generated from PCA [53] can be used. In general, the tensors in the

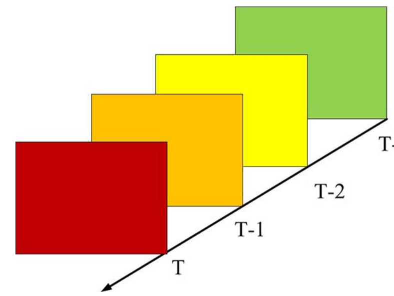

Fig. 1. A 3-order tensor stream: The multimode data are growing incrementally over time.

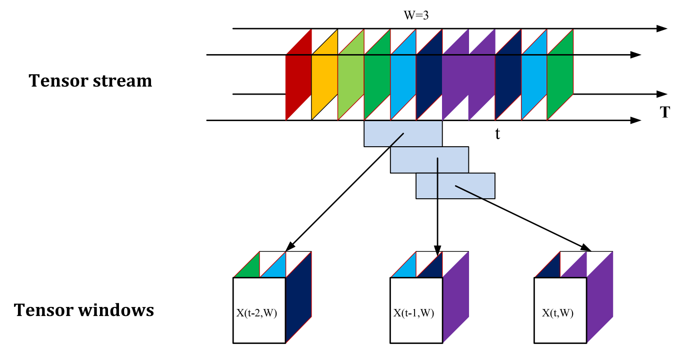

Fig. 2. Tensor windows: The tensor stream can be handle separately by using a tensor window.

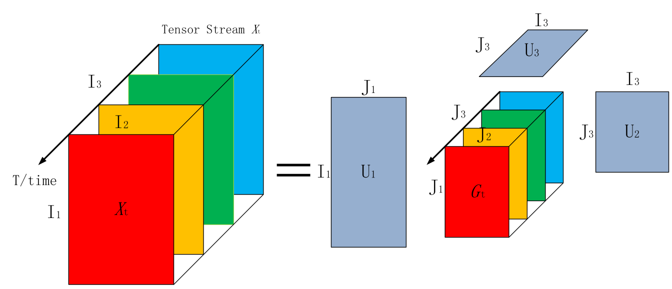

Fig. 3. The decomposition of a 3-order tensor stream.

tensor stream can be explained by M mode factors (basis matrices) $U_m \in R^{I_m \times J_m}$ giving features represented by core tensors $\mathbf{G}_t$. The decomposition of a the tensor stream $\mathbf{X}_t$ can be expressed as:

$$\mathbf{X}_t = \mathbf{G}_t \times_1 U_{(1)} \times_2 U_{(2)}, \ldots, \times_M U_{(M)} + \mathbf{E}_t \tag{6}$$

where the compressed core tensor $\mathbf{G}_t \in R^{J_1 \times J_2, \ldots, \times J_M}$ representing features is of a much lower dimension than the raw data tensor $\mathbf{X}_t$. Each entry of the core tensor $\mathbf{G}_t$ is an individual feature, and expresses the strength of interaction among basis components in different mode factors. Fig. 3 shows an example of decomposition over a sequences of third-order tensor stream.

III. Prediction Method

A. Dynamic Tensor Model for Traffic Flow

The dynamic tensor model for traffic flow is presented in this section. For many spatial-temporal methods, they suppose that the future traffic volume strongly depends on current and past volume of that location and its neighbors [16]. Suppose we need

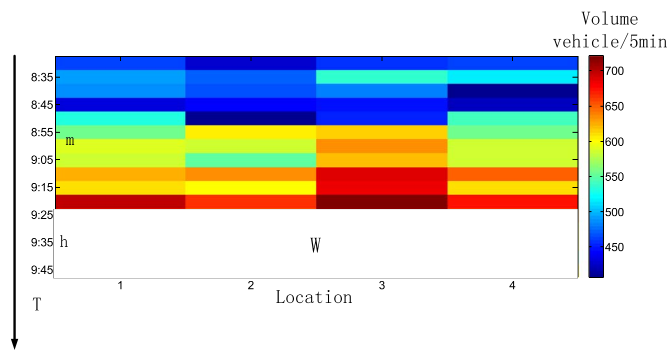

Fig. 4. The dynamic matrix model of the prediction method: The sampling interval is 5 min, the horizon of past value m is 11 (8:30 9:20), and the prediction horizon h is 9 (9:25 9:45).

to forecast traffic volume of p locations $\{s_i\}_{i=1}^{p}$, during the discrete interval $(t_j, t_j+\beta_t, \ldots, t_j+h\beta_t)$. β_t is the sampling interval, h is the prediction horizon. The historical traffic flow data of p locations $\{s_i\}_{i=1}^{p}$, during the $(t_j, t_j-\beta_t, \ldots, t_j-m\beta_t)$ can be used as inputs for generating prediction for traffic flow of $(t_j, t_j+\beta_t, \ldots, t_j+h\beta_t)$. From the view of data modeling, the data model of such prediction methods can be expressed as a dynamic matrix model (see Fig. 4). And the goal of such method methods is to find the relationship function f between current/past traffic data and the future traffic volume such that:

$$Z^W = f(Z^{\overline{W}}) \tag{7}$$

where $W \in R^{p\times h}$, $Z^{\overline{W}} \in R^{p\times m}$. The horizon m determines the horizon of the past value of locations $\{s_i\}_{i=1}^{p}$, which are used for forecasting h steps ahead volume values of $\{s_i\}_{i=1}^{p}$. $\overline{W}$ is the complement set of W.

Nevertheless, the matrix pattern only covers limited multi-mode information of traffic flow data. For example, the matrix pattern in Fig. 4 can represent the traffic flow at some sampling intervals location by location. However, it does not incorporate the traffic data periodic at different sampling intervals, such as day by day and week by week, simultaneously. To tackle this shortcoming of matrix pattern, we construct a 4-way tensor model $\mathbf{Z}_t \in R^{p\times w\times d\times i(m+h)}$ that is incorporated with the day mode information and week mode information from the same time interval of several past weeks. w is the number of historical weeks, d is the number of historical days (a week has 7 days), i is the number of intervals $(m+h)$. Fig. 5 gives an example of such a tensor pattern, and it is clear that the tensor pattern covers the week, day, space and time mode information of traffic flow data.

In real time, the prediction can be implemented from a *traffic flow dynamic tensor window* of width $(m+h)$ on time dimension and a corresponding tensor (see Fig. 2). For any *traffic flow dynamic tensor window*, the traffic data in the prediction horizon can then be predicted based on newly available data, so the prediction problem can be expressed as:

$$\mathbf{Z}_t^{\mathbf{W}_t} = f\left(\mathbf{Z}_t^{\overline{\mathbf{W}}_t}\right) \quad \text{for} \quad t = 1, 2, \ldots, n \tag{8}$$

The prediction problem is then transformed into a dynamic tensor completion problem using Equation (8). And in this paper, the size of dynamic tensor window is selected as following: Some studies [32] suggested that five or six weeks were a suitable amount of historical data for prediction and a number of days greater than six provided no improvement in results. Thus, the week mode dimension of *traffic flow dynamic tensor window* is $w = 5$ or 6. Some research studies [23], [32] suggest to discard the weekend data to run the prediction. However, discarding the weekends would weaken the week mode periodicity. Our goal is to make full use the information from day mode, week mode, location mode and time mode to forecast traffic flow. So the weekends will retain in the dynamic tensor pattern. Day dimension d is set to 7 for our methods, since one week has seven days. It was found in many literatures [12], [13], [23], [32] that a prediction horizon less than 1 h based on 30 min–1 h of available new data obtain a good accuracy for many methods. This choice of horizon is also applied for the proposed dynamic tensor completion method. For example, the historical data width on time dimension m is set to 6 18 and prediction horizon width h is set to 1 12 for a 5-min aggregated traffic flow data.

B. Dynamic Tensor Completion Algorithm

We give the dynamic tensor completion algorithm for short-term traffic flow prediction in this section. A series of studies were carried out in multi-dimension data to find a fast and reliable method that combines and utilizes the multi-mode information for estimating unknown entries of tensor [29], [42], [44], [54]. Tensor-based unknown data estimation methods can capture the global structure of the data via a high-order decomposition. In particular, the most widely used tensor decomposition methods, such as CP decomposition [44] and Tucker decomposition [29], [45], [46], have great ability for multidimensional modeling with many advantages such as fast convergence, high accuracy, and low computational efforts for tensor completion problem [55]. Different from the typical CP and Tucker decomposition, this paper focuses on the decomposition of a dynamic tensor pattern with multi-linear low-n-rank. And many completion methods using the nuclear norm to approximate the n-rank of tensors [42], [49], [54], [56]. In this paper, we use a low-rank matrix factorization model, which has been proved to be faster than many nuclear-norm minimization problems for matrix completion problem [57], to estimate the unknown entries of tensor. The deviation of our algorithm is given as follows:

Suppose that *traffic flow dynamic tensor window* $\mathbf{Z}_t$ has a rank-group of (J_1, J_2, J_3, J_4) for the location, week, day and interval mode with observed entries $\overline{\mathbf{W}}_t$. The deviation starts with the following optimization problem:

$$\min_{Z_{(i)}} \quad \sum_{i}^{4} P_{\overline{W}_{(i)}} \left\| Z_{(i)} - A_i Y_i \right\|_F^2 \tag{9}$$

where $A_i \in R^{m_i\times J_i}, Y_i \in \mathrm{R}^{J_i\times I_i}$; J_i is the mode-n rank of tensor stream $\mathbf{Z}_t$, m_i is the column of $Z_{(i)}$. I_i is the row of $Z_{(i)}$ that equal to the length of the ith dimension of tensor stream $\mathbf{Z}_t$. $P_{\overline{W}_{(i)}}$ is the projection which makes the element out of $\overline{W}_{(i)}$ vanishing.

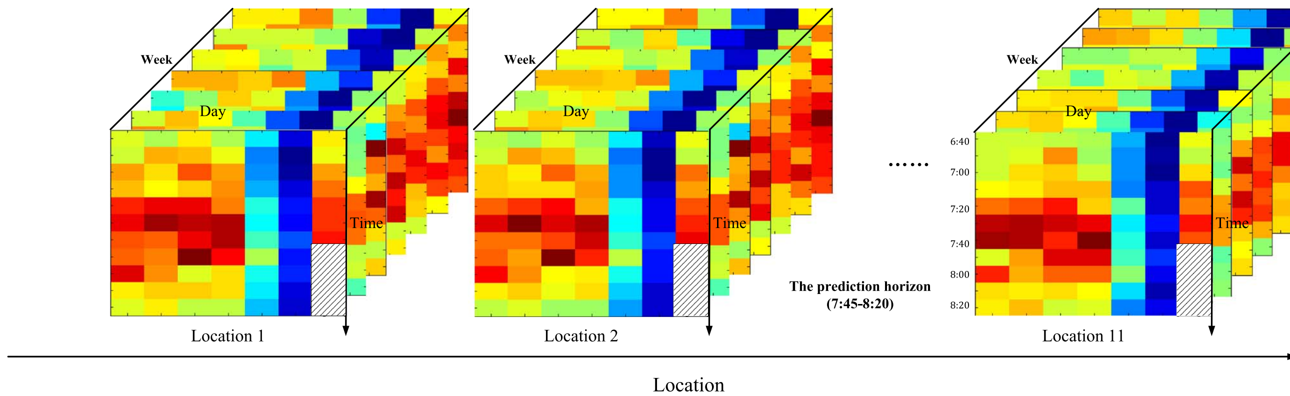

Fig. 5. The dynamic tensor model $\mathbf{Z}_t$ for traffic flow data: The same time periods of past weeks are used to construct the tensor model. (The prediction horizons of each location in this tensor are unknown.)

Because $Z_{(i)}$ $(i = 1, 2, \ldots, N)$ are generated from i-mode unfoldings of tensor $\mathbf{Z}_t$, the problem in Equation (9) is difficult to solve due to the interdependent Frobenius norm constraints of $Z_{(i)}$ $(i = 1, 2, \ldots, N)$. To simplify the formulation, we introduce additional matrices M_1, M_n and obtain the following equivalent formulation (10):

$$\min_{Z_{(i)}, M_i} \quad \sum_{i}^{4} P_{\overline{W}_{(i)}} \|M_i - A_i Y_i\|_F^2$$
$$\text{s:t} \quad P_{\overline{W}_{(i)}}(M_i) = P_{\overline{W}_{(i)}}\left(Z_{(i)}\right). \tag{10}$$

The Frobenius norm are still not independent because of the equality constrain $P_{\overline{W}_{(i)}}(M_i) = P_{\overline{W}_{(i)}}(Z_{(i)})$. Thus we relax the equality constrain by $P_{\overline{W}_{(i)}} \|M_i - Z_{(i)}\|_F^2 \leq d_i$ as Eq (11), so that we can independently solve each subproblem later on

$$\min_{Z_{(i)}, M_i, A_i, Y_i} \quad \sum_{i}^{4} \alpha_i P_{\overline{W}_{(i)}} \|M_i - A_i Y_i\|_F^2 + \beta_i P_{\overline{W}_{(i)}} \left\|M_i - Z_{(i)}\right\|_F^2. \tag{11}$$

Then we propose to employ block coordinate descent (BCD) [58] for the optimization problem in (11), The basic idea of block coordinate descent is to optimize a group of variables while fixing the other groups. We divide the variables into 13 groups: $\mathbf{Z}$, $A_1, A_2, \ldots, A_4$, $Y_1, Y_2, \ldots, Y_4$, $M_1, M_2, \ldots, M_4$. And these variables are solved by each subproblem:

Computing A_i: $A_i \leftarrow M_i Y_i^{\dagger}$
Computing Y_i: $Y_i \leftarrow A_i^{\dagger} M_i$
Computing $\mathbf{Z}$: $\mathbf{Z} \leftarrow \mathbf{Z}^{\mathbf{W}} + P_{W_{(i)}} \left(\sum_i^m \beta_i \text{fold}(M_i) / \sum_i^m \beta_i\right)$
Computing M_i: $T_i = ((\alpha_i A_i Y_i + \beta_i Z_{(i)})/(\alpha_i + \beta_i))$ $M_i \leftarrow T_i + P_{\overline{W}_{(i)}}(Z_{(i)} - T_i)$.

In addition to the optimization problem, another problem is how adaptively choose n-mode rank (J_1, J_2, J_3, J_4) for the prediction model. In the short-term traffic flow prediction, tensor data in each step will be updated dynamically. To exploit the latest information of traffic flow and hence to improve the performance of traffic prediction, it is necessary to adaptively determine the n-mode tensor rank for each updated tensor. Determining the n-mode rank of a tensor is still a challenging problem. Nowadays' approaches such as the DIFFerence in Fit DIFFIT [59], numerical convex hull (NumConvHull) [60] and Automatic relevance determination [61] are computationally expensive and too many models have to be evaluated. In our method, a simple and fast approach, quotient of differences in additional values (QDA) [62], is applied to adaptively determine the n-mode rank. For each mode matricization, the QDA is to find the value of J_m that maximizes,

$$a(J_m) = \frac{\lambda_{J_m} - \lambda_{J_{m+1}}}{\lambda_{J_{m+1}} - \lambda_{J_{m+2}}} \quad \text{for all} \quad \lambda_{J_m} > \text{mean}(\lambda) \tag{12}$$

where $\lambda_{(J_m)}$ is the eigenvalue associated with the J_mth component in a 2-way PCA (eigenvalue decomposition of correlation coefficient matrix) of each mode matricization. It should be noted that the n-mode rank is calculated on the subset $p \times w \times d \times m$ of the dynamic tensor pattern, which means that the complement set contains unknown future value is not used for determining the n-mode rank. Then, the determined n-mode rank is fed to DTC. In this paper, the n-mode rank is re-estimated by QDA every 30 minutes.

For initiation of the algorithm, because the multi-mode property of *traffic flow dynamic tensor windows* between neighboring interval are similar, it is not necessary to randomly initiate $Y_i(t)$ for every *traffic flow dynamic tensor window*. Thus $Y_i(t+1)$ are initiated as $Y_i(t)$ when $t = 2, 3, \ldots, T$. The pseudo-code of the DTC method is given in Algorithm 1.

Algorithm 1: dynamic tensor completion

input : dynamic tensor:$\mathbf{Z}_t$, prediction sets: $\mathbf{W}_t$, rank group $(J_1, J_2, , J_m)$ calculated by QDA, parameter $(\alpha_1, \alpha_2, , \alpha_m)$;$(\beta_1, \beta_2, , \beta_m)$
output: $\mathbf{Z}^W$,A_i,$\mathbf{G}_t = \mathbf{Z}_t \times_1 X_{(1)}^{\mathbf{T}} \times_2 A_{(2)}^{\mathbf{T}} \ldots \times_M A_{(M)}^{\mathbf{T}}$

initialization: set
$Y_i^0 \in \mathrm{R}^{J_i \times I_i}$,$M_i^0 = \textit{unfolding}(P_{\overline{W}_{(i)}}(\mathbf{Z}), i)$
while *not converged* **do**
  $A_i \leftarrow M_i Y_i^{\dagger}$;
  $Y_i \leftarrow X_i^{\dagger} M_i$;
  $\mathbf{Z} \leftarrow \mathbf{Z}^{\mathbf{W}} + P_{W_{(i)}}\left(\frac{\sum_i^m \beta_i fold(M_i)}{\sum_i^m \beta_i}\right)$;
  $M_i \leftarrow T_i + P_{\overline{W}_{(i)}}(Z_{(i)} - T_i)$;
end
$t = t + 1$

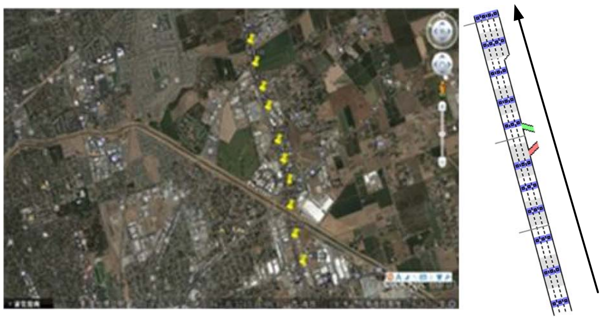

Fig. 6. The locations of this study: These detectors are located on a freeway corridor.

TABLE I
RESULT COMPARISON

Method	Model Type	MAE	MAPE
DTC	4-way tensor(with QDA)	12.9410	9.01%
DTC	4-way tensor(2,2,1,3)	13.1559	9.08%
DTC	3-way tensor(with QDA)	13.0903	9.16%
DTC	3-way tensor(2,1,2)	14.3211	9.15%
Spectral analysis		17.1337	11.41%
ARIMA	p=2,d=1,q=4	18.4157	11.27%

IV. Experiment Results and Analyses

A. The Test Data

To provide a fair comparison, the Performance Measurement System (PeMS) open-access traffic flow datasets [47] are used for this study. The objective of the PeMS project is to collect real-time freeway data from freeways in California and to measure traffic performance.

The particular dataset used in this paper was collected from the adjacent stations located at south bound freeway SR99, District 10, Stanislaus County, California (Fig. 6). The index numbers of these detectors are 1017510, 1017610, 1017710, 1017810, 1017910, 1018110, 1018210, 1018310, 1018410, 1018510, and 1018610. The sampling period is from March 1, 2011 to May 29, 2011. Our analysis is based on the archived 5-minute historical data. Model calibration for some conventional methods is carried out using the data from March 1, 2011 to April 15, 2011, and the traffic data from April 16, 2011 to May 30, 2011 are used for evaluating the prediction performance. The missing data (about 2%) have been imputed by built-in imputation methodology of PeMS. Since the negligible missing ratio, the data set after imputation by PeMS is regarded as an approximately complete data set and used as ground truth in this study.

B. Performance Indexes

In this paper, the mean absolute percentage error (MAPE) is used to compare the performance of traffic forecasting. However, the MAPE will be lower if the traffic volumes are higher. In observance of this, this paper also applies the mean absolute error (MAE) as a complementary measure for MAPE.

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|z_t - N_t| \tag{13}$$

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{|z_t - N_t|}{N_t} \times 100\% \tag{14}$$

where $z_t =$ predicted traffic flow for observation t; $N_t =$ actual traffic flow for observation t; $n =$ number of predictions.

C. Experiment Results

In this section, we compare tensor completion based method with ARIMA [23] (prediction with temporal variabilities), spectral analysis [32] (prediction with temporal variabilities and week mode periodicity). Also, the prediction performance based on 3-way tensor pattern without spatial information and 4-way tensor pattern with global information are compared. Comparing the performance of different forecasting methods is difficult as their objectives can differ considerably from one to another. Since this paper aims at higher prediction accuracy, the performance can be compared roughly under an equivalent condition.

An important problem is the selection of the optimal aggregation time scale for short-term traffic prediction models. Generally, at intervals shorter than 3 min, strong fluctuations are present in traffic flows, which will cause a decline of predictable information [63]; however, the local information within the time interval will also be lost at longer intervals [64]. Considering all factors, the aggregation time scale of traffic volume data is set to 5 min in this paper.

For ARIMA, the Alaike Information Criterion (AIC) is used to measure the goodness of fit for the ARIMA models. And the same day in six different weeks are used as historical days for spectral analysis to run the prediction on that day in the current week. The size of 3-way tensor pattern for dynamic tensor completion (DTC) is $7(\text{weeks}) \times 7(\text{days}) \times 7(\text{intervals})$ on 5 min scale. For the 4-way tensor with spatial information, the size of dynamic tensor pattern is set to $11 \times 7 \times 7 \times 7$. The mode-n-ranks are estimated by QDA. The results without QDA are also given for comparison. The weighted parameters α and β are set as [1, 10, 200, 1] and [100, 100, 100, 100] for 4-way tensor pattern and [10, 200, 1] and [100, 100, 100] for 3-way tensor pattern. The comparison are conducted on a one-step ahead prediction horizon.

All methods are run with Matlab on a Windows Workstation with a Dual-Core Intel(R) Core(TM) 2.50 GHZ CPU and 4 GB RAM. Table I gives the overall results for all the methods.

The results illustrate that the overall results of proposed prediction model DTC is better than matrix-pattern based spectral analysis method and one-way vector-based ARIMA method. The reason is that the proposed DTC method utilizes the information of multiple modes simultaneously, while spectral analysis method only exploits correlations in day and interval modes, and ARIMA only covers interval correlation. Unlike conventional methods, DTC requires neither training process nor a large amount of data for model calibration. For 3-way tensor model, only $7 \times 7 \times 7 - 1 = 342$ historical data points are sufficient to perform reliable prediction.

From Table I, it is easy to see that DTC with and/or without QDA outperforms traditional methods, and DTC with QDA is better than DTC with empirical mode-n-rank. For DTC without

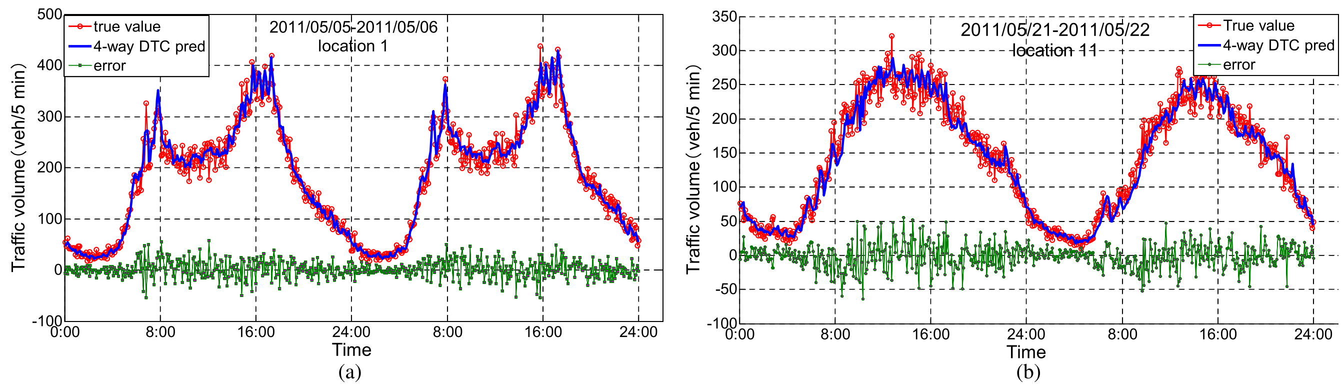

Fig. 7. The prediction results of 4-way DTC. (a) Workdays. (b) Weekends.

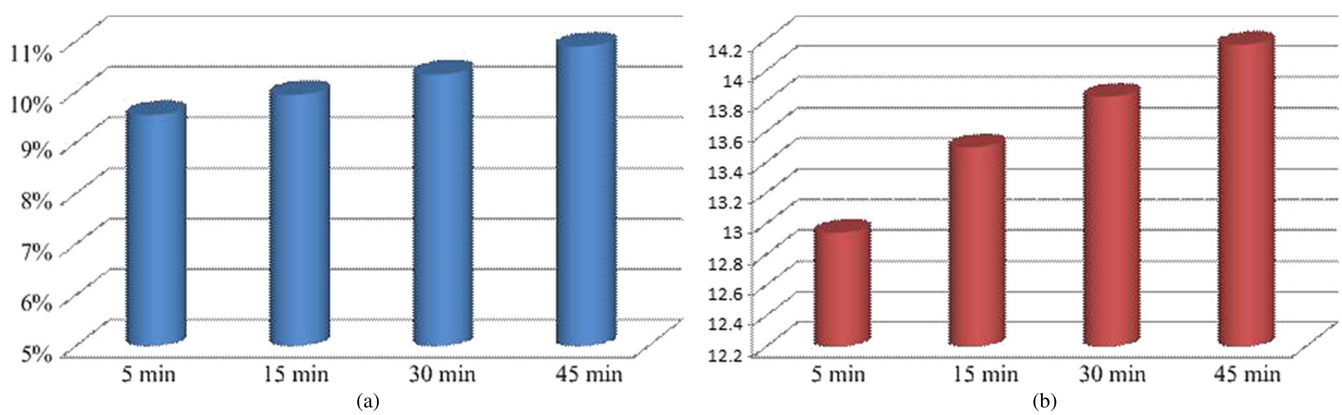

Fig. 8. Prediction results of different horizons: The longer horizon does not lower down much of the prediction accuracy. (a) MAPEs. (b) MAEs.

QDA, all possible n-rank from (2, 1, 1) to (4, 4, 4) for 3-way tensor and from (2, 1, 1, 1) to (4, 4, 4, 4) for 4-way tensor have been investigated, and n-rank (2, 1, 2) and (2, 2, 1, 3) listed in table achieves the best performance. This observation indicates that QDA can help effectively capture the time-varying structure of a dynamic tensor, and hence leads to a better performance.

Fig. 7 gives parts prediction results and prediction errors of 4-way DTC on workdays and weekends. As shown in Fig. 7, DTC exhibits reliable prediction results on both workdays and weekends. The overall accuracy of night traffic flow is higher than that of daytime traffic flow. Nevertheless, as the same as other prediction methods, the accuracy of DTC slightly declines when traffic flow is sharply changed, which will be a focus of our future research.

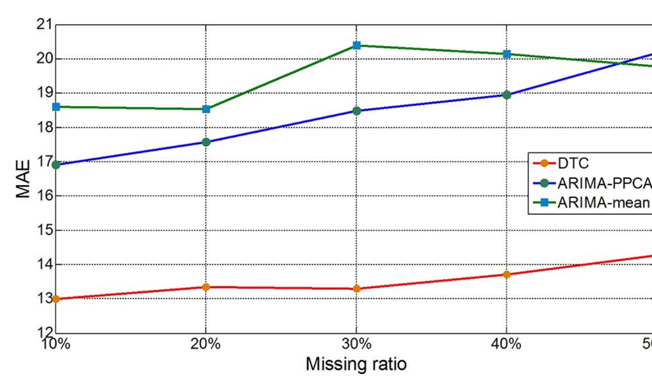

Fig. 9. The MAE curves of different methods under various missing ratios.

V. Further Discussion

The previous experimental comparisons and conclusions were drawn based on the predictions of the one-step ahead horizon prediction and all experiment are made under complete dataset. In this section, we give some further discussion including prediction with longer horizon, prediction under missing data and the computation complexity about DTC to show the superiority of DTC.

A. Longer Horizon Prediction

Predicting traffic flow on several time horizons in the future allows for a wider range of applications to make use of the predictions. For DTC, predicting a longer horizon only means that more entries of real-time tensor pattern are set to unknown. Simply, for a prediction horizon of 15 min with 5 min scale data, the first 6 data points are used to forecast the 7–9th data points (horizon of 15 min) and the 7–9th data points are set to unknown in the dynamic tensor pattern. Then, the 9-period tensor window of size $11 \times 7 \times 7 \times 9$ incorporates the 7–9th data points and the model is refit to forecast the 10–12th data points. The MAPEs and MAEs of such a dynamic tensor with horizons of 5 min (1-step), 15 min (3-step), 30 min (6-step) and 45 min (9-step) are given in Fig. 8.

It can be found that the longer horizon did not deteriorate the prediction accuracy significantly. Even when the prediction

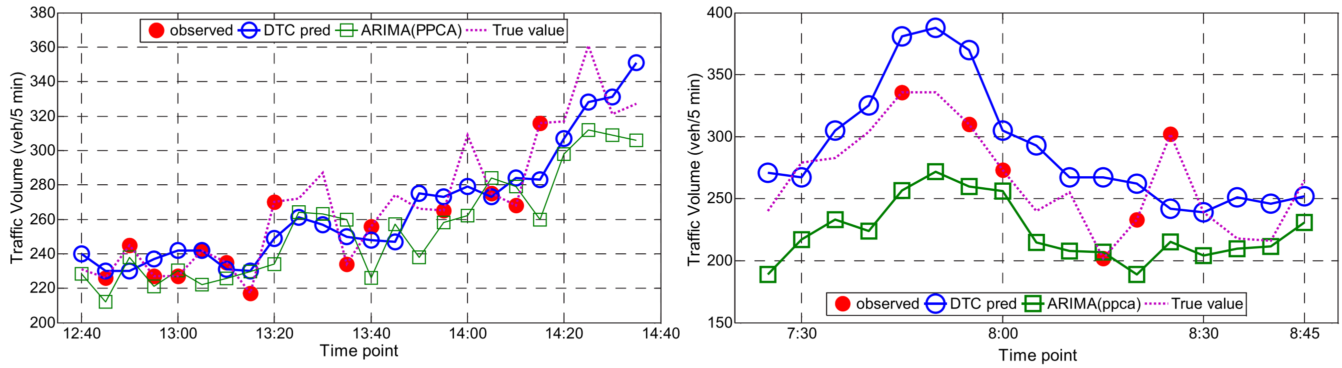

Fig. 10. A comparison between DTC and ARIMA(PPCA) under 50% missing data. Two examples from noon and morning peak hours. It can be found that DTC is more reliable with missing data of high ratio.

horizon reaches 45 min (9-step ahead), the MAPE and MAE of DTC is 10.61% and 14.1113 respectively. It is not much bigger than that of 9.01% and 12.9410 for 1-step horizon ahead prediction. It implies that the performance of the DTC is not closely related to the prediction horizon. The reason is that the tensor completion methods, which makes full use of the global information of the multi-mode data, can still accurately estimate the unknown data even when large number of entries are unavailable.

B. *Prediction Under Missing Data*

As discussed in [23], [24], [65],short-term traffic prediction on incomplete data is a challenge with strong and very urgent practicality. Different from traditional prediction models that processing missing data imputation and traffic flow prediction in different framework, dynamic tensor completion can impute missing data and forecast traffic flow simultaneously, which means that DTC can forecast traffic flow without pre-processing of missing traffic data. For DTC, the missing data contained within tensor pattern is set to unknown together with the data within prediction horizon, and then estimated by dynamic tensor completion in each step.

In the experiment, DTC is compared with ARIMA where the missing data are imputed by online-PPCA(a matrix data pattern based missing data imputation approach) [66] and ARIMA model that use historical mean data to substitute missing data [65] (our method is different from the method in [65], we use mean of same interval traffic flow rather than the mean of all the traffic flow). The total missing ratio is set to a number in [10%, 50%] to better test the relationship of prediction performance and missing ratio. since every random missing case gives different result, ten times of experiments are conducted for a single missing ratio, and then the average MAPEs and MAEs for different ratios are given. And in the experiment, both the historical datasets and test datasets contain missing data. We also test an extreme missing data case where several days' traffic data of a one location are all missing. In the extreme case, the traffic data from April 18, 2011 to April 24, 2011 for location 1 are set to missing for testing. The results of random case are given in Figs. 9 and 10.

From Fig. 9, it can be found that the MAE curve of DTC rises gradually as the missing ratio increases and DTC achieves

TABLE II
THE MAEs OF DTC AND ARIMA(MEAN) FOR CONSECUTIVELY MISSING DAYS

Missing days	DTC	ARIMA(mean substitution)
1	16.2629	16.5977
2	16.6002	17.5962
3	16.8700	18.4311
4	17.0585	20.2589
5	17.5774	21.9772
6	18.0714	26.0432
7	18.5003	26.0432

the lowest error in all missing ratio. Whereas the two ARIMA models deteriorate when missing ratio is higher than 30%. This is reasonable due to the fact that tensor pattern is more robust under high missing ratio. Fig. 10 gives a comparison between under 50% missing data. It can be found that DTC is more reliable with missing data of high ratio.

For extreme case, the missing data is set from 1 to 7 days. Table II shows the MAPEs and MAEs of DTC within the missing period (April 18, 2011 to April 24, 2011) for location 7. and Fig. 11 shows the performance results against raw data when all traffic data from April 18, 2011 to April 24, 2011 are all missing. Notice that the matrix-based method (PPCA) cannot fulfill the work when one or several days of traffic volume data are missing. The underlying reason is related to the coherence in matrix decomposition, namely, a matrix can never be recovered if there are the whole missing row (or a column) of a matrix [67]. Since the extreme missing case is corresponding to one or several rows missing in matrix pattern, which makes online-PPCA fail to handle missing data, the results of ARIMA method handle missing data by online-PPCA is not given in this paper.

As shown in Table II and Fig. 11, the traffic data under missing data within 7 consecutive days can be accurately predicted by the proposed method. The errors are growing slowly with the number of missing days. Even when 7 days' traffic data are all missing, the most prediction results are still reliable except some abnormal pattern of weekends. The main reason is that DTC can automatically use others' mode information to estimate traffic data when one or two mode information are lost. Under this case several day mode information and time

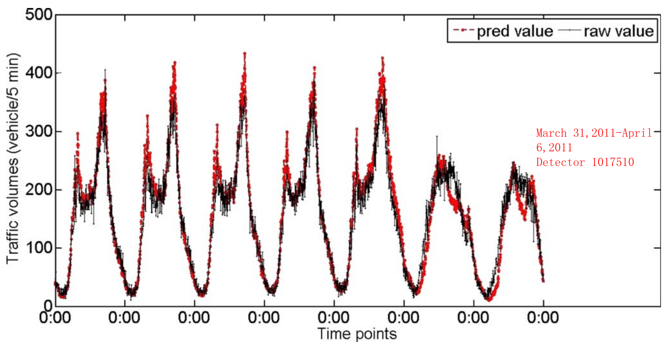

Fig. 11. The prediction results of the missing periods (the 7 days' traffic data are all missing).

mode information are lost when 7 consecutive days' traffic data are missing. But DTC can still successfully work with week mode and space mode information of traffic data. In contrast, the matrix pattern (PPCA) only utilizes time mode and day mode information fails to handle missing data under this extreme case.

C. Computation Complexity of DTC Approach

One of the advantages of DTC over many conventional models is that dynamic tensor pattern can capture multi-mode information effectively on a smaller data set. Unlike conventional methods, DTC requires neither training process nor a large amount of data for model calibration. The results show that $7 \times 7 \times 7 - 1$ historical data is sufficient for the proposed method to perform reliable prediction in each step. The running time of DTC promises real-time prediction. The experiments mentioned before were running on a 64-bit machine with 4-GB memory and MATLAB 2012b environment. Average running time using DTC (spatial-temporal tensor pattern with QDA) for each step is about 0.6 seconds, with 500 maximal iteration and $1e - 5$ tolerance. The reason is that the tensor pattern is dramatically reduced by the dynamic tensor approach, and the proposed tensor completion approach is based on a fast matrix factorization approach.

The rank-group (J_1, J_2, J_m) also controls the computation complexity of DTC. In out works, the rank-group are determined by QDA in each step. We test many fixed rank-groups for DTC. It is found that the rank of [2, 2, 1, 3] achieve the best performance for the 4-way dynamic tensor of size $11 \times 7 \times 7 \times 7$. It indicates that the traffic flow data has a low-n-rank property over the space, day, week and time mode. We also had investigated different datasets from the PeMS. The results are similar and DTC also provides reliable results.

VI. Conclusion

In this paper, we proposed a novel data prediction method, namely Dynamic Tensor Completion (DTC). The proposed DTC makes an effective use of multi-mode periodicity such as daily and weekly periodicity, spatial information as well as temporal variations of traffic flow by representing traffic information as dynamic tensor pattern and estimating future traffic under the low-n-rank assumption. Experiments on benchmark datasets show that DTC outperforms conventional vector-based and matrix-based prediction approaches in terms of prediction accuracy. Besides, another important advantage is that DTC is able to accurately forecast traffic flow with missing data. This makes DTC useful for practical ITS applications, since that real data collection always suffer from unreliability and dysfunction, which easily make the collected data incomplete. Moreover, the experimental results show that the speed of DTC meets the online prediction requirement. For future work, our approach might be improved by considering non-linear factorization of tensor based traffic data. The potential difficulty of scalability of our method lies in how to adapt our method on traffic flow prediction of transportation network. In such case, the size of tensor and the amount of missing data in dynamic tensor will be very large and the tensor completion method proposed in this paper might not be suited. The combination of tensor completion and recently proposed tensor network concepts [68] may provide a potential solution. Also, we think that giving the spatial information such as the distance between locations and the number of lanes is very important, and it is our future work.

Acknowledgment

The authors would like to thank B. Liu of China University of Petroleum (East China) for his suggestive discussions.

References

[1] J. Zhang *et al.*, "Data-driven intelligent transportation systems: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1624–1639, 2011.
[2] B. Ran, P. J. Jin, D. Boyce, T. Z. Qiu, and Y. Cheng, "Perspectives on future transportation research: Impact of intelligent transportation system technologies on next-generation transportation modeling," *J. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 226–242, 2012.
[3] J. Wang, L. Zhang, D. Zhang, and K. Li, "An adaptive longitudinal driving assistance system based on driver characteristics," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 1, pp. 1–12, Mar. 2013.
[4] J. Wang, S. Li, Y. Zheng, and X. Lu, "Longitudinal collision mitigation via coordinated braking of multiple vehicles using model predictive control," *Integr. Comput.-Aided Eng.*, vol. 22, no. 2, pp. 171–185, Apr. 2015.
[5] J. Wang, J. Wu, and Y. Li, "The driving safety field based on driver–vehicle–road interactions," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 14, pp. 2203–2214, Aug. 2015.
[6] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-term traffic forecasting: Where we are and where were going," *Transp. Res. C, Emerging Technol.*, vol. 43, pp. 3–19, Jun. 2014.
[7] E. I. Vlahogianni, J. C. Golias, and M. G. Karlaftis, "Short-term traffic forecasting: Overview of objectives and methods," *Transp. Rev.*, vol. 24, no. 5, pp. 533–557, 2004.
[8] M. Van Der Voort, M. Dougherty, and S. Watson, "Combining Kohonen maps with ARIMA time series models to forecast traffic flow," *Transp. Res. C, Emerging Technol.*, vol. 4, no. 5, pp. 307–318, 1996.
[9] C. Han, S. Song, and C. H. Wang, "A real-time short-term traffic flow adaptive forecasting method based on ARIMA model," *Acta Simulata Systematica Sinica*, vol. 7, p. 043, 2004.
[10] Z. Li, Y. Li, and L. Li, "A comparison of detrending models and multi-regime models for traffic flow prediction," *IEEE Intell. Transp. Syst. Mag.*, vol. 6, no. 4, pp. 34–44, Winter 2014.
[11] B. L. Smith, B. M. Williams, and R. Keith Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transp. Res. C, Emerging Technol.*, vol. 10, no. 4, pp. 303–321, Aug. 2002.
[12] 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 Syst. Appl.*, vol. 36, no. 3, pp. 6164–6173, Apr. 2009.
[13] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Optimized and meta-optimized neural networks for short-term traffic flow prediction: A

genetic approach," *Transp. Res. C, Emerging Technol.*, vol. 13, no. 3, pp. 211–234, Jun. 2005.

[14] X. Jiang and H. Adeli, "Dynamic wavelet neural network model for traffic flow forecasting," *J. Transp. Eng.*, vol. 131, no. 10, pp. 771–779, 2005.

[15] S. Ishak, P. Kotha, and C. Alecsandru, "Optimization of dynamic neural network performance for short-term traffic prediction," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1836, no. 1, pp. 45–56, 2003.

[16] W. Min and L. Wynter, "Real-time road traffic prediction with spatio-temporal correlations," *Transp. Res. C, Emerging Technol.*, vol. 19, no. 4, pp. 606–616, 2011.

[17] S. Dunne and B. Ghosh, "Regime-based short-term multivariate traffic condition forecasting algorithm," *J. Transp. Eng.*, vol. 138, no. 4, pp. 455–466, 2011.

[18] A. Stathopoulos and M. G. Karlaftis, "A multivariate state space approach for urban traffic flow modeling and prediction," *Transp. Res. C, Emerging Technol.*, vol. 11, no. 2, pp. 121–135, 2003.

[19] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Spatio-temporal short-term urban traffic volume forecasting using genetically optimized modular networks," *Comput.-Aided Civil Infrastructure Eng.*, vol. 22, no. 5, pp. 317–325, 2007.

[20] L. Li, Y. Li, and Z. Li, "Efficient missing data imputing for traffic flow by considering temporal and spatial dependence," *Transp. Res. C, Emerging Technol.*, vol. 34, pp. 108–120, Sep. 2013.

[21] Y. Li, Z. Li, and L. Li, "Missing traffic data: Comparison of imputation methods," *IET Intell. Transp. Syst.*, vol. 8, no. 1, pp. 51–57, 2014.

[22] H. Tan, Y. Wu, B. Cheng, W. Wang, and B. Ran, "Robust missing traffic flow imputation considering nonnegativity and road capacity," *Math. Problems Eng.*, vol. 2014, pp. 1–8, 2014.

[23] C. Chen, Y. Wang, L. Li, J. Hu, and Z. Zhang, "The retrieval of intra-day trend and its influence on traffic prediction," *Transp. Res. C, Emerging Technol.*, vol. 22, pp. 103–118, Jun. 2012.

[24] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results," *J. Transp. Eng.*, vol. 129, no. 6, pp. 664–672, 2003.

[25] W. C. Hong, "Application of seasonal SVR with chaotic immune algorithm in traffic flow forecasting," *Neural Comput. Appl.*, vol. 21, no. 3, pp. 583–593, 2012.

[26] L. Li, X. Su, Y. Zhang, Y. Lin, and Z. Li, "Trend modeling for traffic time series analysis: An integrated study," *IEEE Intell. Transp. Syst. Mag.*, vol. 16, no. 6, pp. 3430–3439, Dec. 2015.

[27] Y. Zhang, Y. Zhang, and A. Haghani, "A hybrid short-term traffic flow forecasting method based on spectral analysis and statistical volatility model," *Transp. Res. C, Emerging Technol.*, vol. 43, pp. 65–78, Jun. 2014.

[28] M. C. Tan, S. C. Wong, J. M. Xu, Z. R. Guan, and P. Zhang, "An aggregation approach to short-term traffic flow prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 1, pp. 60–69, Mar. 2009.

[29] H. Tan *et al.*, "A tensor-based method for missing traffic data completion," *Transp. Res. C, Emerging Technol.*, vol. 28, pp. 15–27, Mar. 2013.

[30] D. Ni, J. D. Leonard, A. Guin, and C. Feng, "Multiple imputation scheme for overcoming the missing values and variability issues in ITS data," *J. Transp. Eng.*, vol. 131, no. 12, pp. 931–938, 2005.

[31] M. T. Asif, N. Mitrovic, L. Garg, J. Dauwels, and P. Jaillet, "Low-dimensional models for missing data imputation in road networks," in *Proc. IEEE ICASSP*, 2013, pp. 3527–3531.

[32] T. T. Tchrakian, B. Basu, and M. O'Mahony, "Real-time traffic flow forecasting using spectral analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 519–526, Jun. 2012.

[33] A. Cichocki, R. Zdunek, A. H. Phan, and S. I. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*. Chichester, U.K.: Wiley, 2009.

[34] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.

[35] E. Acar and B. Yener, "Unsupervised multiway data analysis: A literature survey," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 1, pp. 6–20, Jan. 2009.

[36] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear subspace analysis of image ensembles. In computer vision and pattern recognition," in *Proc. IEEE Comput. Soc. Conf.*, 2003, vol. 2, pp. 2–93.

[37] E. Acar, D. M. Dunlavy, and T. G. Kolda, "Link prediction on evolving data using matrix and tensor factorizations," in *Proc. IEEE ICDMW*, 2009, pp. 262–269.

[38] M. Li and V. Monga, "Robust video hashing via multilinear subspace projections," *IEEE Trans. Image Process.*, vol. 21, no. 10, pp. 4397–4409, Oct. 2012.

[39] H. Lee, Y. D. Kim, A. Cichocki, and S. Choi, "Nonnegative tensor factorization for continuous EEG classification," *Int. J. Neural Syst.*, vol. 17, no. 4, pp. 305–317, 2007.

[40] H. Tan, J. Feng, G. Feng, W. Wang, and Y. J. Zhang, "Traffic volume data outlier recovery via tensor model," *Math. Prob. Eng.*, vol. 2013, pp. 1–8, 2013.

[41] J. Dauwels *et al.*, "Predicting traffic speed in urban transportation subnetworks for multiple horizons," in *Proc. IEEE 13th ICARCV*, Dec. 2014, pp. 1–6.

[42] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.

[43] Y. Xu, R. Hao, W. Yin, and Z. Su, "Parallel matrix factorization for low-rank tensor completion," *Inverse Problems & Imaging*, vol. 9, no. 2, 2015.

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

[45] H. Tan, J. Feng, Z. Chen, F. Yang, and W. Wang, "Low multilinear rank approximation of tensors and application in missing traffic data," *Adv. Mech. Eng.*, vol. 6, Jan.–Dec. 2014, Art. ID 157597.

[46] H. Tan, B. Cheng, J. Feng, L. Liu, and W. Wang, "Mixture augmented Lagrange multiplier method for tensor recovery and its applications," *Discrete Dyn. Nat. Soc.*, vol. 2014, pp. 1–9, 2014, Art. ID 914963.

[47] California Highway Patrol (CHP), 2006. [Online]. Available: http://www.chp.ca.gov/index.html

[48] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1996.

[49] S. Gandy, B. Recht, and I. Yamada, "Tensor completion and low-n-rank tensor recovery via convex optimization," *Inverse Problems*, vol. 27, no. 2, 2011, Art. ID 025010.

[50] H. Tan, B. Cheng, W. Wang, Y. J. Zhang, and B. Ran, "Tensor completion via a multi-linear low-n-rank factorization model," *Neurocomputing*, vol. 133, pp. 161–169, 2014.

[51] J. Sun, D. Tao, S. Papadimitriou, P. S. Yu, and C. Faloutsos, "Incremental tensor analysis: Theory and applications," *ACM Trans. Knowl. Discovery Data*, vol. 2, no. 3, p. 11, 2008.

[52] L. Shi, A. Gangopadhyay, and V. P. Janeja, "STenSr: Spatio-temporal tensor streams for anomaly detection and pattern discovery," *Knowl. Inf. Syst.*, pp. 1–21, 2014.

[53] A. H. Phan and A. Cichocki, "Tensor decompositions for feature extraction and classification of high dimensional datasets," *IEICE Nonlinear Theory Appl.*, vol. 1, no. 1, pp. 37–68, 2010.

[54] M. Signoretto, R. Van de Plas, B. De Moor, and J. A. Suykens, "Tensor versus matrix completion: A comparison with application to spectral data," *IEEE Signal Process. Lett.*, vol. 18, no. 7, pp. 403–406, Jul. 2011.

[55] M. Ishteva, "Numerical methods for the best low multilinear rank approximation of higher-order tensors," Ph.D. dissertation, Dept. Elect. Eng., Katholieke Universiteit Leuven, Leuven, Belgium, 2009.

[56] R. Tomioka, T. Suzuki, K. Hayashi, and H. Kashima, "Statistical performance of convex tensor decomposition," in *Proc. NIPS*, 2011, pp. 972–980.

[57] Z. Wen, W. Yin, and Y. Zhang, "Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm," *Math. Program. Comput.*, vol. 4, no. 4, pp. 333–361, 2012.

[58] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *J. Optim. Theory Appl.*, vol. 109, no. 3, pp. 475–494, 2001.

[59] M. E. Timmerman and H. A. Kiers, "Three-mode principal components analysis: Choosing the numbers of components and sensitivity to local optima," *Brit. J. Math. Statist. Psychol.*, vol. 53, no. 1, pp. 1–16, 2000.

[60] E. Ceulemans and H. A. Kiers, "Selecting among three-mode principal component models of different types and complexities: A numerical convex hull based method," *Brit. J. Math. Statist. Psychol.*, vol. 59, no. 1, pp. 133–150, 2006.

[61] M. Mørup and L. K. Hansen, "Automatic relevance determination for multi-way models," *J. Chemometrics*, vol. 23, no. 7/8, pp. 352–363, 2009.

[62] J. Niesing, *Simultaneous Component and Factor Analysis Methods for Two or More Groups: A Comparative Study*. Leiden, The Netherlands: DSWO Press, Leiden Univ.. 1997.

[63] C. Oh, S. G. Ritchie, and J. S. Oh, "Exploring the relationship between data aggregation and predictability to provide better predictive traffic information," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1935, no. 1, pp. 28–36, 2005.

[64] E. Vlahogianni and M. Karlaftis, "Temporal aggregation in traffic data: Implications for statistical characteristics and model choice," *Transp. Lett.*, vol. 3, no. 1, pp. 37–49, 2011.
[65] H. Chen, S. Grant-Muller, L. Mussone, and F. Montgomery, "A study of hybrid neural network approaches and the effects of missing data on traffic forecasting," *Neural Comput. Appl.*, vol. 10, no. 3, pp. 277–286, Dec. 2001.
[66] L. Qu, L. Li, Y. Zhang, and J. Hu, "PPCA-based missing data imputation for traffic flow volume: A systematical approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 3, pp. 512–522, Sep. 2009.
[67] E. J. Cands and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.
[68] A. Cichocki, "Era of big data processing: A new approach via tensor networks and tensor decompositions," arXiv:1403.2048, 2009, to be published.

Huachun Tan (M'13) received the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2006. He is an Associate Professor with the School of Mechanical Engineering, Beijing Institute of Technology, Beijing. His research interests include image engineering, pattern recognition, and intelligent transportation systems.

Yuankai Wu received the Bachelor's degree from the Department of Mechanical Engineering, Shanghai Ocean University, Shanghai, China, in 2012 and the Master's degree from the Department of Transportation Engineering, Beijing Institute of Technology, Beijing, China, in 2015. He is currently working toward the Ph.D. degree with the Department of Transportation Engineering, Beijing Institute of Technology. His research interests include intelligent transportation systems and machine learning.

Bin Shen (S'09–M'14) received the B.S. and M.S. degrees from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2007 and 2009, respectively, and the M.S. and Ph.D. degrees from the Department of Computer Science, Purdue University, West Lafayette, IN, USA, in 2011 and 2014, respectively. He is currently with Google Research, New York, NY, USA. His research interests include image processing, machine learning, and data mining.

Peter J. Jin received the B.S. degree in automation from the Department of Automation, College of Information Technology, Tsinghua University, Beijing, China, in 2009, and the M.S. and Ph.D. degrees in civil and environmental engineering from University of Wisconsin–Madison, Madison, WI, USA, in 2007 and 2009, respectively. He is an Assistant Professor with the Department of Civil and Environmental Engineering and an affiliated member of the Center for Advanced Infrastructure and Transportation at Rutgers State University of New Jersey, Piscataway, NJ, USA. His research area is intelligent transportation systems, with special interest in mobile sensor and content data, connected vehicles, and active transportation and demand management. His recent research interests include transportation big data analytics, vehicle-to-infrastructure communications, and unmanned airborne vehicle application in traffic monitoring and incident management.

Bin Ran received the B.S. degree in civil engineering from Tsinghua University, Beijing, China, in 1986; the M.S. degree in civil engineering from the University of Tokyo, Tokyo, Japan, in 1989; and the Ph.D. degree in civil engineering from the University of Illinois, Chicago, IL, USA, in 1993. He is a Professor with the Department of Civil and Environmental Engineering, University of Wisconsin–Madison, Madison, WI, USA. He has led the development and deployment of various traffic information systems and technologies in the United States and China. He is the author of two leading textbooks on dynamic traffic networks. He has coauthored more than 100 journal papers and more than 200 referenced papers at national and international conferences. His research has focused on five major areas: 1) intelligent transportation system; 2) dynamic transportation network and traffic modeling; 3) development of cellular probe technologies; 4) connected vehicle, and 5) Internet of mobility.