Ride the Flow: Dynamic Tensor and Adaptive Modeling for Short-Term Traffic Prediction

Yilin Liu[1] and Guillaume Sartoretti[2]

***Abstract*— Accurate short-term traffic flow prediction plays a pivotal role in enabling reliable and intelligent decision-making for automated transportation systems. However, the inherent complexity of spatiotemporal traffic patterns and the frequent presence of missing data present critical challenges to prediction accuracy and system resilience. In this paper, we present a dynamic tensor-based forecasting framework tailored for intelligent transportation scenarios. We first employ a hierarchical clustering strategy to identify key intersections with strong interdependencies, enhancing the modeling of spatial correlations. Subsequently, we introduce a dynamic "Location–X–Time" tensor structure that captures multi-scale temporal dependencies and adapts in real-time through a sliding window mechanism, ensuring robustness to data incompleteness. Our approach improves the reliability of downstream applications such as adaptive traffic control and vehicle-infrastructure coordination. Experimental results on large-scale real-world datasets demonstrate that our proposed method consistently outperforms state-of-the-art models, supporting the advancement of secure and intelligent urban mobility systems.**

I. Introduction

Accurate traffic flow prediction is a fundamental component of intelligent transportation systems (ITS), supporting real-time traffic management, resource allocation, and vehicle-infrastructure coordination [1]. With the growing adoption of connected vehicles, roadside sensing units, and mobile computing platforms, vast streams of spatiotemporal traffic data are continuously generated across urban road networks [2], [3]. Moreover, traffic flow prediction can potentially support intelligent applications such as coordinated traffic signal control [4], [5]. Vehicular Social Networks (VSNs) have emerged as a key paradigm in this context, enabling vehicles and infrastructure to cooperatively sense, communicate, and share mobility information [6]. While this collaborative environment enhances situational awareness and system responsiveness, it also introduces critical challenges for predictive modeling—particularly the high dimensionality, dynamic variability, and frequent incompleteness of traffic data.

Conventional traffic prediction methods, ranging from temporal and spatial models [7]–[9] to matrix decomposition [10] and tensor-based techniques [11]–[13] have demonstrated improved forecasting performance. However, they often struggle to model strong interdependencies between road segments, adapt to multi-scale temporal dynamics, or maintain robustness under data loss conditions. More recently, machine learning-based approaches, such as neural networks [14], [15], support vector machines (SVM) [16], clustering models [17], and hybrid deep learning frameworks [18]–[20], have pushed predictive accuracy further. Nonetheless, these methods generally assume high-quality inputs, limiting their practical use in complex, real-world traffic environments where missing or noisy data is common [21], [22].

To improve adaptability and generalizability, recent works have explored diverse learning frameworks, including adaptive KNN models [23], fuzzy neural networks [24], [25], and recurrent architectures such as LSTMs [26]. Others turned to network theory [27] or probabilistic reasoning [28] to better capture traffic dynamics. Hybrid approaches, such as clustering with genetic algorithms [29] or pattern mining for ferry systems [30] have shown promise, but often overlook the full potential of high-dimensional spatiotemporal data. In this context, tensor decomposition techniques may be able to model these multi-way structures more effectively [31].

Despite recent advances, key challenges remain in intelligent traffic systems and vehicle-infrastructure cooperation: (1) Limited spatial modeling – most methods overlook fine-grained correlations between adjacent intersections, reducing regional consistency; (2) Underuse of high-dimensional structures - simplified 1D/2D representations may miss complex spatiotemporal interactions; (3) Poor robustness to missing data – sensor failures and transmission issues often cause data gaps, undermining prediction reliability.

In response, this paper proposes a dynamic tensor-based framework for short-term traffic flow forecasting, designed to support more robust, data-driven decision-making in smart transportation networks. The method includes:

1) **A hierarchical clustering algorithm**, which identifies spatially correlated intersections, enhancing the relevance and accuracy of traffic flow prediction;
2) **A tensor-based prediction model** with a flexible "Location–X–Time" structure that can capture multi-level temporal dynamics and exploits the multidimensional nature of traffic data;
3) **A sliding window mechanism**, which enables dynamic tensor transformation, improving prediction robustness and accuracy under missing data scenarios.

Our proposed framework improves spatial modeling and robustness to missing data, enabling accurate traffic forecasting aligned with key ITS goals. The rest of the paper

This work is supported by A*STAR, CISCO Systems (USA) Pte. Ltd and National University of Singapore under its Cisco-NUS Accelerated Digital Economy Corporate Laboratory (Award I21001E0002).

[1]Yilin Liu is with the State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications. liuyilin10@bupt.edu.cn

[2]Guillaume Sartoretti is with the Department of Mechanical Engineering, National University of Singapore. guillaume.sartoretti@nus.edu.sg

is organized as follows: Section II reviews the background; Section III presents our model; Section IV reports our experimental results; Finally, Section V concludes with key findings and future work.

II. Problem Formalization

A. Related Definitions

Definition 1 (Traffic Flow Time Series)**.** The traffic flow time series at a given intersection i is defined as $v_i = \{x_{i1}, x_{i2}, x_{i3}, \ldots, x_{in}\}$, where x_{in} represents the traffic flow at intersection i during the nth short time period. The set v_i denotes the complete time series of traffic flow at intersection i over a continuous time span.

Definition 2 (Tensor Stream)**.** A tensor stream is defined as a sequence of N*th-order* tensors $\{T_1, T_2, \ldots, T_{t_{\max}}\}$, where each tensor $T_t \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$. The sequence is considered as an $(N+1)$*th-order* tensor stream when $t_{\max}$ increases dynamically over time [9].

Definition 3 (Tensor Window). $W(t, w) \in \mathbb{R}^{w \times I_1 \times \cdots \times I_N} = (T_{t-w+1}, \ldots, T_t)$ is a sliding window that captures a finite sub-sequence of the tensor stream at time t with size w. The tensor window enables localized analysis of temporal traffic variations, facilitating robust short-term prediction.

Definition 4 ("Location-X-Time" Tensor Model)**.** The tensor representation of traffic flow data is denoted as $A \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, where I_1 represents the location dimension (set of intersections), I_2 represents the variable parameter X (multi-level temporal representation), I_3 represents the time dimension (discrete time intervals).

B. Problem Formulation

We consider m intersections within a given spatial region, each recording a time series of traffic flow $\{v_1, v_2, \ldots, v_m\}$, which forms a traffic flow matrix $V_{m \times n} = [v_1 \; v_2 \; \ldots \; v_m]^T$. To extract meaningful correlations, a hierarchical clustering algorithm is applied to group intersections with similar traffic flow patterns, resulting in k high-similarity traffic clusters $\{v_1, v_2, \ldots, v_k\}, k \le m$.

We then construct a tensor-based prediction model using the "Location-X-Time" framework, where X is a variable parameter controlling multi-level temporal relationships. The traffic tensor is defined as $A \in \mathbb{R}^{I_1 \times I_2 \times I_3}$. To improve adaptability, we introduce a sliding tensor window $W(t, w)$ to perform dynamic tensor decomposition for real-time prediction. The predicted traffic tensor $\hat{A}_t$ is computed following:

$$\left.\begin{matrix} V_{m \times n} \\ A \end{matrix}\right\} \rightarrow W(t, w) \rightarrow \hat{A}_t$$

Our model addresses two key tasks: intersection correlation analysis, and short-term traffic prediction. In this work, we use the Euclidean distance threshold p to measure traffic similarity. Traffic flows with distances below p are clustered via hierarchical clustering $cluster(v_i \mid \textit{Euclidean distance} < p)$. Our tensor model, built on clustered traffic data, can then integrate a sliding window $\hat{A}_t \in R^{I_1 \times I_2 \times I_3}$ for dynamic tensor decomposition, enabling short-term traffic prediction.

III. Proposed Model and Algorithms

To address the above issues, we propose a short-term traffic flow prediction method. Traffic data is modeled as a dynamic tensor to preserve its temporal and multi-dimensional structure. The approach consists of three stages: (i) intersection correlation analysis based on traffic sequence similarity; (ii) dynamic tensor modeling, and (iii) short-term prediction using sliding window and tensor decomposition techniques (see Fig. 1).

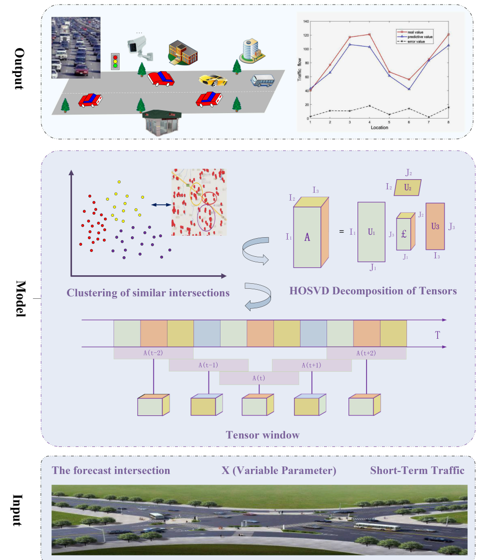


Fig. 1. Framework for short-term traffic flow prediction.

A. Intersection Correlation Analysis

Traffic flows exhibit both temporal dependencies and spatial correlations across intersections. To address this, we first apply agglomerative hierarchical clustering to group intersections with similar traffic dynamics. Each intersection's time series $v_i = \{x_{i1}, x_{i2}, \ldots, x_{in}\}$ is treated as an individual cluster. A pairwise distance matrix $D_{m \times m}$ is defined as:

$$D_{m \times m} = \begin{bmatrix} 0 & d_{12} & \ldots & d_{1m} \\ d_{12} & 0 & \ldots & d_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m1} & d_{m2} & \ldots & 0 \end{bmatrix}, \quad (1)$$

where the Euclidean distance between time series v_i and v_j is computed as:

$$d_{ij} = \sqrt{\sum_{k=1}^{n} (x_{ik} - x_{jk})^2}. \quad (2)$$

The algorithm iteratively merges the most similar clusters until the predefined number of clusters q is reached. This ensures that intersections with correlated traffic behaviors are grouped together, strengthening spatial correlations and enhancing the robustness of short-term traffic flow prediction.

B. Dynamic Modeling

We then propose a dynamic tensor decomposition model to predict traffic flow, leveraging data compression and mapping to uncover correlations and handle missing values.

1) Tensor Modeling Based On Variable Parameter: Tensor modeling for traffic flow data utilizes a third-order tensor $A \in R^{I_1 \times I_2 \times I_3}$ as defined in Definition 4. The first order (tensor dimension) represents predicted intersections, the second order corresponds to the variable parameter X, and the third order captures time values for predictions. The variable parameter X accounts for temporal characteristics, enabling analysis at daily and weekly levels. The tensor element values C are defined as:

$$C = \begin{cases} \textit{Traffic flow} & \text{short-term traffic flow} \\ 0 & \text{others} \end{cases}$$

The values of short-term traffic flows fluctuate over time. We select 10 intersections within the target zone, with short-term intervals set between 5 and 15 minutes. When traffic flow data is unrecorded or missing, it is represented as 0 or an alternative estimated value.

2) Tensor Decomposition Method: The core idea of tensor decomposition is to uncover dimensional correlations through data compression and mapping. Iterative reconstruction restores the tensor, facilitating similarity measurement for short-term flow prediction. This paper develops a third-order tensor model based on the "Location-X-Time" structure, formulated as:

$$A \approx \pounds \times_1 U_1 \times_2 U_2 \times_3 U_3, \tag{3}$$

where A is decomposed into 3 factor matrices U_1, U_2, U_3, and a core tensor $\pounds$. Our aim is to solve the minimization problem $\min \|A - \pounds \times_1 U_1 \times_2 U_2 \times_3 U_3\|$.

We apply HOSVD (Higher-Order SVD) to a 3-order traffic flow tensor A, and we define the three matrices unfolding operations as follows $A_{(1)} \in R^{I_1 \times (I_2 I_3)}$, $A_{(2)} \in R^{I_2 \times (I_1 I_3)}$, $A_{(3)} \in R^{I_3 \times (I_1 I_2)}$, where $A_{(1)}$, where $A_{(2)}$ and $A_{(3)}$ are called the 1-, 2- and 3-mode matrix unfolding of A as shown in Fig. 2.

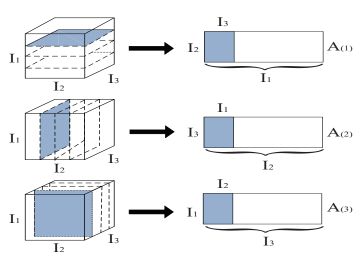


Fig. 2. Each mode matrix of tensor.

Then, we apply singular value decomposition (SVD) to each $A_{(n)}, 1 \le n \le 3$. The SVD decomposition of the 1-mode matrix unfolding is:

$$A_{(1)} = U_{I_1 \times I_1} \cdot \Sigma_{I_1 \times I_2 I_3} \cdot V^T_{I_2 I_3 \times I_2 I_3}. \tag{4}$$

The characteristic matrix Σ is a diagonal matrix arranged according to the singular values of matrix $A_{(1)}$, where a greater singular value indicates greater importance of the feature to the matrix $A_{(1)}$. The matrix $A_{(1)}$ is usually described by its m_1 ($m_1 < \min\{I_1, I_2 I_3\}$) most important features. After m_1 is determined, we use the left singular matrix U and the right singular matrix V^T to mine the latent information under implicit attention, resulting in the reconstruction matrix $\hat{A}_{(1)}$

$$\hat{A}_{(1)} = U_{I_1 \times m_1} \cdot \Sigma_{m_1 \times m_1} \cdot V^T_{m_1 \times I_2 I_3} = U^{(1)} \cdot \Sigma^{(1)} \cdot V^{T^{(1)}}. \tag{5}$$

Similarly, we can obtain $\hat{A}_{(2)}$ and $\hat{A}_{(3)}$. The core tensor based on the left singular matrix finally reads

$$\pounds = A \times_1 U^{(1)^T} \times_2 U^{(2)^T} \times_3 U^{(3)^T}. \tag{6}$$

We finally reconstruct the traffic flow tensor as:

$$\hat{A} = \pounds \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}. \tag{7}$$

The elements in $\hat{A}$ represent short-term traffic flow. By applying tensor decomposition, data compression, and mapping procedures, we obtain an approximate tensor for prediction. Given a dynamic traffic flow tensor window A_t with rank-group (r_1, r_2, r_3) for location, X, time modes, and observed entries $W_{(i)}$:

$$\min_{A_{(i)}} \sum_i^3 P_{W_{(i)}} \left\| A_{(i)} - X_i Y_i \right\|_F^2, \tag{8}$$

where $X_i \in R^{m_i \times r_i}$, $Y_i \in R^{r_i \times n_i}$; r_i is the mode-n rank, m_i and n_i are the respective dimensions of tensor stream A_t, and $P_{W_{(i)}}$ ensures projection constraints. For simplicity, we introduce auxiliary matrices $M_1, M_2, \ldots, M_n$ and reformulate:

$$\min_{A_{(i)}, M_i} \sum_i^3 P_{W_{(i)}} \left\| M_i - X_i Y_i \right\|_F^2. \tag{9}$$

Since the Frobenius norm constraints remain dependent due to $P_{W_{(i)}}(M_i) = P_{W_{(i)}}(A_{(i)})$, we relax them:

$$\min_{A_{(i)}, M_i, X_i, Y_i} \sum_i^3 \alpha_i P_{W_{(i)}} \left\| M_i - X_i Y_i \right\|_F^2 + \beta_i P_{W_{(i)}} \left\| M_i - X_i Y_i \right\|_F^2. \tag{10}$$

We use block coordinate descent, where each variable is optimized one at a time while keeping the others fixed, gradually approaching the optimal solution

- $X_i \leftarrow M_i Y_i^{\dagger}$, where $\dagger$ denotes the pseudo-inverse.
- $Y_i \leftarrow X_i^{\dagger} M_i$.

The optimized pattern matrix retains key features. SVD decomposition extracts essential information, and tensor decomposition transforms multi-dimensional data, enabling accurate short-term traffic flow prediction.

3) Missing Value Processing: In traffic datasets, missing data occurs when no traffic information is recorded for an intersection at a specific time. Formally, this means the tuple (I_1, I_2, I_3) lacks a valid value. Factors such as weather, road conditions, and equipment failures contribute to these gaps, leading to incomplete records in the system.

If there is a non-zero value in the traffic data, it must satisfy $A = A_m + A_\Omega$, where A_m represents the set of missing value elements, and A_Ω represents the set of non-zero elements, which satisfy the following formula:

$$\min_{\hat{A}_t} \left\| A - \hat{A}_t \right\| = \sum\nolimits_{(I_1 I_2 I_3 \in \Omega)} (A - \hat{A}_t). \tag{11}$$

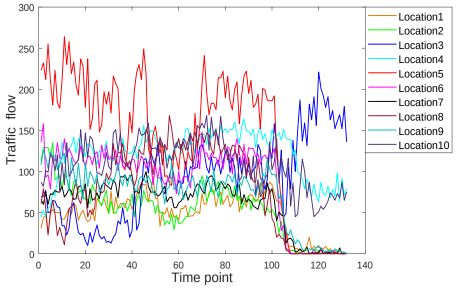
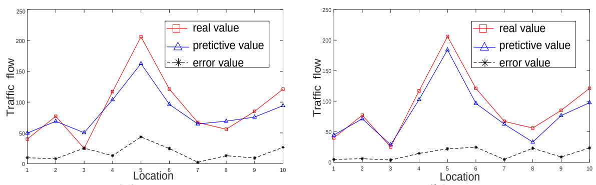


Fig. 3. Traffic flow distribution at 10 intersections.



(a) (b)

Fig. 4. Ten intersections prediction results with variable parameter X. (a) X = Day. (b) X = Week.

Common methods for handling missing values include setting them to zero, averaging the available data, or using the tensor decomposition method of Eq. 3 to iteratively estimate the missing values $A_m^1, A_m^2, \ldots$. The process satisfies the following condition:

$$\min_{A_m^{t+1}} \left\| \left(A_\Omega + A_m^t\right) - A^{t+1} \right\|, \tag{12}$$

where A_m^t is the solution of the t-th iteration of the missing value part in A^t, A^{t+1} represents the (t+1)-th iteration value. When the difference between two successive iterations satisfies the convergence condition, the iteration is terminated. The iterative values A^{t+1} obtained at this time are close to the tensor $A_{I_1 I_2 I_3}$. Since missing values from previous iterations are utilized in each iteration, they are gradually optimized throughout the calculation process, ultimately enabling a more accurate short-term traffic flow prediction.

4) Iterative Algorithm: The input of our algorithm includes the time series matrix of traffic flow $V_{m\times n}$ and traffic flow data tensor A $\in R^{I_1\times I_2\times I_3}$, as explained in Section 3, as well as the classification threshold p, mode rank (r_1, r_2, r_3) and parameters $(\alpha_1, \alpha_2, \alpha_3)$, $(\beta_1, \beta_2, \beta_3)$. As shown in Algorithm 1, our iterative algorithm can get the similarity traffic flow v_i after clustering and prediction results $\hat{A}_t$ by tensor decomposition.

The time complexity of calculating traffic flow time series similarity is $O(n^3)$. The time complexity of the coordinate descent method to update parameters is $O(n)$. When short-term traffic flow is calculated based on tensor decomposition, the resource cost is concentrated primarily in the process of decomposition, and the complexity is $O(n_{I_1} * n_{I_2} * n_{I_3})$, which is related to the size of the tensor. The overall time complexity for our algorithm is $O(n_{I_1} * n_{I_2} * n_{I_3}) + O(n^3) + O(n) \sim O(n_{I_1} * n_{I_2} * n_{I_3}) + O(n^3)$.

IV. Experiments and Analysis

A. Experimental Settings

This section describes our experimental data, baselines, and evaluation metrics used to assess the proposed approach.

Algorithm 1 Iterative Algorithm.

Input:
Traffic flow matrix at time t: $V_{m\times n} = [v_1\ v_2\ \ldots\ v_m]^T$
Traffic flow tensor $A \in \mathbb{R}^{I_1\times I_2\times I_3}$ with missing values
Parameters: p, $(\alpha_1, \alpha_2, \alpha_3)$, $(\beta_1, \beta_2, \beta_3)$, rank (r_1, r_2, r_3)
Output:
Similar $pass\_port$ v_i (intersection number)
Approximate tensor $\hat{A}_t$
// Initialization
Initialize: $t = 0$, Y_i, M_i
// Calculate distances
Calculate d_{ij} according to Eq. (2)
Calculate all pair-wise distances of $v_1, v_2, \ldots, v_m$ to other clusters less than p
Construct tensor and perform tensor decomposition:
    Matrix unfolding: $A_{(1)}, A_{(2)}, A_{(3)}$
Calculate $A^t = A_m^t + A_\Omega$
while not reaching max-iteration **do**
    Update $X_i \leftarrow M_i Y_i^\dagger$ (Eq. 12)
    Update $Y_i \leftarrow X_i^\dagger M_i$ (Eq. 12)
    Perform matrix decomposition on $A_{(1)}, A_{(2)}, A_{(3)}$ (Eq. 4)
    Approximate matrices $\hat{A}_{(1)}, \hat{A}_{(2)}, \hat{A}_{(3)}$ (Eqs. 5–7)
    Calculate core tensor $\pounds$ (Eq. 8)
    Reconstruct approximate tensor $\hat{A}_t$ (Eq. 9)
end while
// Train parameters and predict
$t \leftarrow t + 1$
repeat
    Construct tensor A_{t+1} based on A_t
    Reconstruct approximate tensor A_{t+1}
until convergence

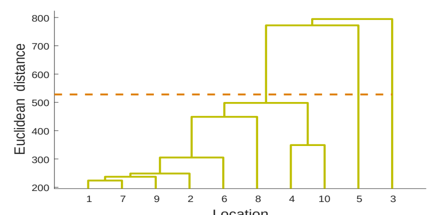
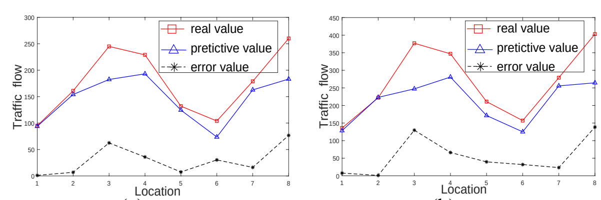


Fig. 5. Hierarchical clustering of traffic flow time series.

1) Experimental Data: We utilize real-world vehicle traffic data collected from multiple intersections in Chongqing, China. Each record includes vehicle license plates, timestamps, vehicle locations, and other attributes. The dataset contains millions of daily records, with key identifiers such as plate number ($license\_no$), timestamp ($pass\_time$), intersection name ($port\_name$), intersection number ($pass\_port$), and geographic coordinates (lng, lat). A sample of the data is presented in Table. I.

We focus on 10 key intersections with high traffic flow, recorded at 5-minute intervals from 8:00 to 19:00 each day. Fig. 3 illustrates the fluctuations in traffic flow across intersections. Understanding these variations is crucial for analyzing correlation patterns and optimizing traffic predic-



(a) (b)

Fig. 6. Results for eight intersections with extended intervals. (a) 10min. (b) 15min.

TABLE I Data sample.

license_no	*pass_time*	*pass_port*	*port_name*	*lat*	*lng*
A	7:09	1.41E+1	Port 1	318.340	1172.494
B	4:39	1.41E+2	Port 2	318.341	1172.268
C	0:05	1.41E+3	Port 3	318.342	1172.288
D	9:51	1.41E+4	Port 4	318.343	1172.797

tion models.

2) Baseline Methods: To evaluate the performance of our method, we compare it to the following baseline methods.

- **SCDT (Similarity Clustering Dynamic Tensor)**: A dynamic tensor model incorporating traffic similarity clustering for short-term flow prediction.
- **ARIMA** [32]: A linear time-series model leveraging historical traffic data to forecast short-term traffic flow.
- **SVR** [33]: A regression approach mapping traffic data to high-dimensional space for better accuracy.
- **BPNN** [34]: A neural network model trained on historical traffic data, adjusting weights and thresholds to minimize prediction error.

3) Evaluation Metrics: We evaluate the predictive performance using mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE):

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n}\left|x_t - M_t\right|, \tag{13}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{x_t - M_t}{x_t}\right| \times 100\%, \tag{14}$$

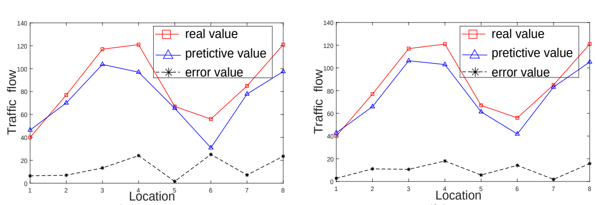
$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(x_t - M_t\right)^2}, \tag{15}$$

where x_t represents the predicted traffic flow, M_t is the actual traffic flow, and n is the number of predictions. Lower values indicate higher prediction accuracy.

B. Prediction Performance Analysis

We analyze performance by exploring various tensor parameter settings, clustering strategies, and handling different data intervals and sparsity levels.

1) Comparison of Different Parameter Settings: We first examine how the parameter X (daily vs. weekly) affects prediction accuracy. The intersection dimension is set to 10, the time dimension is 18. Then, we compare two configurations of daily (10 days) and weekly (8 weeks) in Fig. 4.



(a) (b)

Fig. 7. Eight intersections prediction results with variable parameter X. (a) X = Day, sampling interval = 5min. (b) X = Week, sampling interval = 5min.

When X is set to Week, performance shows a slight improvement over day, suggesting that weekly patterns capture more consistent traffic behavior.

2) Clustering Division Comparison: To further improve accuracy, hierarchical clustering is used to group intersections with similar traffic patterns, as shown in Fig. 5.

We compare prediction performance before and after clustering in Table II. To examine the impact of short time periods on forecasting accuracy, we extend the duration of the short time period appropriately. Specifically, we select sampling intervals of 10 minutes and 15 minutes and predict short-term traffic flow for the time periods 9:50–10:00 and 9:45–10:00 for the last day of the data set and calculate the average. The experimental results are presented in Fig. 6. Clustering yields more accurate predictions, and weekly settings again outperform daily.

TABLE II Prediction performance before & after clustering.

Similarity Clustering	Parameter X	MAE	MAPE	RMSE
Before (10 intersections)	Day	17.61	20.91%	21.12
Before (10 intersections)	Week	13.45	15.40%	15.92
After (8 intersections)	Day	13.49	16.41%	16.12
After (8 intersections)	Week	**10.01**	**11.81%**	**11.49**

3) Impact of Different Time Intervals: We assess prediction performance across sampling intervals (5, 10, and 15 minutes), as shown in Table III. Longer intervals reduce accuracy due to increased traffic uncertainty. Fig. 7 shows that when focusing on eight clustered intersections, the proposed model performs better under the weekly dimension. Longer sampling intervals generally lead to higher error.

4) Comparison with Baseline Algorithms: Fig. 8 and Table. IV show that the proposed method closely matches real traffic flow and outperforms BPNN, SVR, and ARIMA in both accuracy and robustness. Unlike baseline models that require retraining, it delivers fast, reliable predictions (within 0.8136s) and adapts well to weekly patterns. Performance remains stable even with 10–50% missing data, demonstrating strong resilience. Fig. 9 shows that prediction error increases with higher ratios of missing data, but our SCDT method maintains lower MAE and MAPE than other models, showing strong robustness to data sparsity.

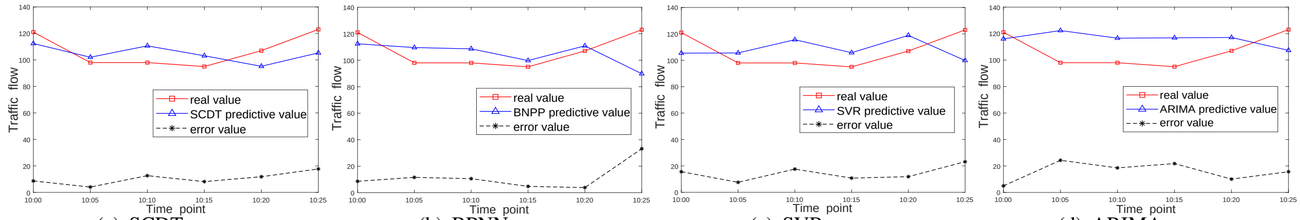
V. Conclusion

We propose a dynamic tensor-based framework for short-term traffic prediction that leverages hierarchical clustering and adaptive modeling to capture spatial-temporal patterns and handle missing data. Experiments on real-world datasets demonstrate superior accuracy and robustness, with future work focusing on incorporating contextual information like V2X to improve generalizability.

References

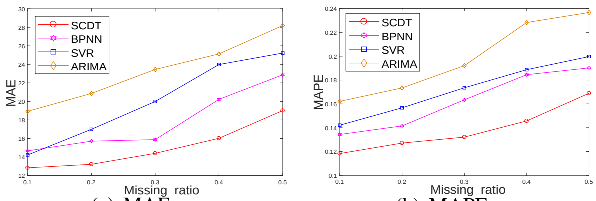
[1] X. Kong, M. Li, T. Tang, K. Tian, L. Moreira-Matias, F. Xia, Shared subway shuttle bus route planning based on transport data analytics, IEEE Trans. Autom. Sci. Eng. 15 (4) (2018) 1507–1520.

[2] X. Kong, F. Xia, J. Wang, A. Rahim, S. K. Das, Time-location-relationship combined service recommendation based on taxi trajectory data, IEEE Trans. Ind. Inform. 13 (3) (2017) 1202–1212.



(a) SCDT (b) BPNN (c) SVR (d) ARIMA

Fig. 8. Prediction results of each algorithm.



(a) MAE (b) MAPE

Fig. 9. Impact of missing data on prediction performance. (a) MAE vs. missing ratio. (b) MAPE vs. missing ratio.

TABLE III Performance across different time intervals.

Time Interval	X	MAE	MAPE	RMSE
5min	Week	**10.01**	**11.81%**	**11.49**
10min	Week	29.65	15.01%	39.32
15min	Week	54.76	17.68%	73.82

[3] B. J. Yanling Cui, F. Zhang, Highway traffic condition detection based on data fusion, Acta computer science 40 (8) (2017) 1789–1812.

[4] Y. Zhang, H. Goel, P. Li, M. Damani, S. Chinchali, G. Sartoretti, Coordlight: Learning decentralized coordination for network-wide traffic signal control, IEEE Trans. Intell. Transp. Syst. (2025).

[5] Y. Zhang, Y. Liu, P. Gong, P. Li, M. Fan, G. Sartoretti, Unicorn: A universal and collaborative reinforcement learning approach towards generalizable network-wide traffic signal control, arXiv preprint arXiv:2503.11488 (2025).

[6] X. Kong, F. Xia, Z. Ning, A. Rahim, Y. Cai, Z. Gao, J. Ma, Mobility dataset generation for vehicular social networks based on floating car data, IEEE Trans. Veh. Technol. 67 (5) (2018) 3874–3886.

[7] Z. H. Mir, F. Filali, An adaptive kalman filter based traffic prediction algorithm for urban road network, in: Proc. 2016 12th Int. Conf. Innov. Inf. Technol. (IIT), IEEE, 2016, pp. 1–6.

[8] W. Hu, L. Yan, K. Liu, H. Wang, Pso-svr: A hybrid short-term traffic flow forecasting method, in: Proc. 2015 IEEE 21st Int. Conf. Parallel Distrib. Syst. (ICPADS), IEEE, 2015, pp. 553–561.

[9] J. Qu, X. Gu, L. Zhang, Improved ugrnn for short-term traffic flow prediction with multi-feature sequence inputs, in: 2018 Int. Conf. Inf. Netw. (ICOIN), IEEE, 2018, pp. 13–17.

[10] H. Tan, Y. Wu, B. Shen, P. J. Jin, B. Ran, Short-term traffic prediction based on dynamic tensor completion, IEEE Trans. Intell. Transp. Syst. 17 (8) (2016) 2123–2133.

[11] H. Tan, Y. Wu, G. Feng, W. Wang, B. Ran, A new traffic prediction method based on dynamic tensor completion, Procedia-Social and Behavioral Sciences 96 (2013) 2431–2442.

[12] T. Xie, S. Li, L. Fang, L. Liu, Tensor completion via nonlocal low-rank regularization, IEEE Trans. Cybern. 49 (6) (2018) 2344–2354.

[13] Y. Wu, H. Tan, Y. Li, J. Zhang, X. Chen, A fused cp factorization method for incomplete tensors, IEEE Trans. Neural Netw. Learn. Syst. 30 (3) (2018) 751–764.

[14] L. Zhang, Q. Liu, W. Yang, N. Wei, D. Dong, An improved k-nearest neighbor model for short-term traffic flow prediction, Procedia-Social and Behavioral Sciences 96 (2013) 653–662.

[15] N. Polson, V. Sokolov, Bayesian particle tracking of traffic flows, IEEE Trans. Intell. Transp. Syst. 19 (2) (2017) 345–356.

[16] D. Chen, Research on traffic flow prediction in the big data environment based on the improved rbf neural network, IEEE Transactions on Industrial Informatics 13 (4) (2017) 2000–2008.

[17] H. Jiang, Y. Zou, S. Zhang, J. Tang, Y. Wang, Short-term speed prediction using remote microwave sensor data: Machine learning versus statistical model, Math. Probl. Eng. 2016 (1) (2016) 9236156.

TABLE IV Evaluation metrics for each algorithm.

Algorithm	MAE	MAPE	RMSE	Mean Pred. Time	Training Time
SCDT	10.48	9.68%	11.31	0.81	No training
BPNN	13.66	10.45%	15.59	0.02	18.25
SVR	13.42	13.30%	15.29	0.35	196.17
ARIMA	**15.89**	**15.49%**	**17.23**	**0.13**	**19.35**

[18] M. Duo, Y. Qi, G. Lina, E. Xu, A short-term traffic flow prediction model based on emd and gpso-svm, in: 2017 IEEE 2nd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC), IEEE, 2017, pp. 2554–2558.

[19] S. Du, T. Li, X. Gong, Y. Yang, S. J. Horng, Traffic flow forecasting based on hybrid deep learning framework, in: 2017 Int. Conf. Intell. Syst. Knowl. Eng.(ISKE), IEEE, 2017, pp. 1–6.

[20] H. Yang, Y. Zou, Z. Wang, B. Wu, A hybrid method for short-term freeway travel time prediction based on wavelet neural network and markov chain, Can. J. Civ. Eng. 45 (2) (2018) 77–86.

[21] S. Wang, G. Mao, Missing data estimation for traffic volume by searching an optimum closed cut in urban networks, IEEE Trans. Intell. Transp. Syst. 20 (1) (2018) 75–86.

[22] G. Chi, L. Jingnan, F. Yuan, L. Meng, C. Jingsong, Value extraction and collaborative mining method of location big data [j], Journal of Software 25 (004) (2014) 713–730.

[23] B. Sun, W. Cheng, P. Goswami, G. Bai, Short-term traffic forecasting using self-adjusting k-nearest neighbours, IET Intelligent Transport Systems 12 (1) (2018) 41–48.

[24] J. Tang, F. Liu, Y. Zou, W. Zhang, Y. Wang, An improved fuzzy neural network for traffic speed prediction considering periodic characteristic, IEEE Trans. Intell. Transp. Syst. 18 (9) (2017) 2340–2350.

[25] J. Tang, F. Liu, W. Zhang, R. Ke, Y. Zou, Lane-changes prediction based on adaptive fuzzy neural network, Expert systems with applications 91 (2018) 452–463.

[26] Q. Zhuo, Q. Li, H. Yan, Y. Qi, Long short-term memory neural network for network traffic prediction, in: 2017 12th Int. Conf. Intell. Syst. Knowl. Eng. (ISKE), IEEE, 2017, pp. 1–6.

[27] Y. Li, L. Zhao, Z. Yu, S. Wang, Traffic flow prediction with big data: A learning approach based on sis-complex networks, in: 2017 IEEE 2nd Inf. Technol., Netw., Electron. Autom. Control Conf. (ITNEC), IEEE, 2017, pp. 550–554.

[28] Y. Zou, J. E. Ash, B.-J. Park, D. Lord, L. Wu, Empirical bayes estimates of finite mixture of negative binomial regression models and its application to highway safety, Journal of Applied Statistics 45 (9) (2018) 1652–1669.

[29] J. Tang, G. Zhang, Y. Wang, H. Wang, F. Liu, A hybrid approach to integrate fuzzy c-means based imputation method with genetic algorithm for missing traffic volume data estimation, Transportation Research Part C: Emerging Technologies 51 (2015) 29–40.

[30] W. Zhang, J. Tang, H. Kristian, Y. Zou, Y. Wang, Hybrid short-term prediction of traffic volume at ferry terminal based on data fusion, IET Intelligent Transport Systems 10 (8) (2016) 524–534.

[31] W. Shao, L. Chen, License plate recognition data-based traffic volume estimation using collaborative tensor decomposition, IEEE Trans. Intell. Transp. Syst. 19 (11) (2018) 3439–3448.

[32] C. Xu, Z. Li, W. Wang, Short-term traffic flow prediction using a methodology based on autoregressive integrated moving average and genetic programming, Transport 31 (3) (2016) 343–358.

[33] J. Ahn, E. Ko, E. Y. Kim, Highway traffic flow prediction using support vector regression and bayesian classifier, in: 2016 Int. Conf. Big Data Smart Comput. (BigComp), IEEE, 2016, pp. 239–244.

[34] C. Yanchong, H. Darong, Z. Ling, A short-term traffic flow prediction method based on wavelet analysis and neural network, in: 2016 Chinese Control and Decision Conference (CCDC), IEEE, 2016, pp. 7030–1034.