Adaptive Truncated Schatten Norm for Traffic Data Imputation with Complex Missing Patterns

Haopeng Deng, Fucheng Zheng, Kaixiang Ma, Xinhai Xia*

School of Intelligent Transportation and Engineering (School of Future Transportation), Guangzhou Maritime University, Guangzhou, 510725, China; E-Mail: xiaxinhai@126.com

ABSTRACT

Accurate traffic data imputation is essential for Intelligent Transportation Systems (ITS), particularly under complex missing patterns. This study presents the Adaptive and Truncated Schatten Norm Low-Rank Tensor Completion (LRTC-ATSN) model, which employs the flexible Schatten norm for low-rank approximation and introduces an adaptive truncation mechanism to suppress noise and retain core features. To address non-convex optimization, the model integrates the Adan algorithm with Nesterov momentum, achieving equilibrium between computational efficiency and recovery precision through dynamic parameter tuning. To evaluate performance, a framework was designed to simulate diverse real-world traffic scenarios with mixed missing patterns. Extensive experiments on datasets from Guangzhou and Seattle demonstrate that LRTC-ATSN outperforms existing methods, yielding 10.6% lower MAPE and 6.1% lower RMSE relative to the best baseline model. Even with 95.85% data loss, the model maintains high reliability. These results underscore LRTC-ATSN's potential for enhancing ITS data robustness and applications across domains like finance and healthcare.

keywords: Traffic Data Imputation, Low-Rank Tensor Completion, Missing Data Patterns, Adaptive Update, Non-Convex Optimization, Truncated Schatten Norm

INTRODUCTION

Intelligent Transportation Systems (ITS) are integral to modern urban traffic management, ensuring both efficiency and safety through optimizing traffic flow, route planning, and signal control. However, as L. Yang et al. (2024) highlight, missing traffic data—especially during peak hours or on critical road segments—can impair system performance, leading to inaccuracies in traffic regulation and suboptimal signal control. These challenges make traffic data imputation a vital research priority.

Traffic data imputation methods are typically divided into three groups: temporal, spatial, and spatiotemporal. Temporal methods, such as the Laplacian Convolution Representation (LCR) model (X. Chen, Cheng, et al., 2024), focus on capturing time dependencies for imputation tasks. Spatial methods, like the RDGCNI model (Y. Chen et al., 2022), analyze traffic state propagation across stations utilizing dynamic graph

convolution networks. While effective, temporal methods often overlook spatial dependencies, whereas spatial methods fail to capture temporal dynamics. These have led to growing interest in spatiotemporal methods to achieve more robust imputation.

Low-Rank Tensor Completion (LRTC) has shown promise for traffic data imputation by capturing spatiotemporal patterns. The classical HaLRTC model (Liu et al., 2012), using the ADMM algorithm to minimize the tensor nuclear norm, treats all singular values equally, limiting its imputation accuracy. To mitigate this, the Truncated Nuclear Norm (TNN) method (X. Chen, J. Yang, et al., 2020) retains dominant singular values while regularizing the smaller ones. Building on this, Nie et al. (2022) proposed the LRTC-TSpN model, which uses the truncated Schatten norm to further enhance recovery accuracy, but lacks sufficient parameter flexibility across scenarios.

Despite these advancements, LRTC models still struggle with handling complex missing data patterns and the substantial computational burden they impose. To break these limitations, we propose the Adaptive and Truncated Schatten Norm Low-Rank Tensor Completion model (LRTC-ATSN). By integrating the Adan algorithm with Nesterov momentum (Xie et al., 2024), LRTC-ATSN accelerates parameter convergence while maintaining high imputation precision. Moreover, in contrast to existing methods tailored to specific missing patterns, LRTC-ATSN is capable of accommodating mixed missing patterns, better reflecting real-world traffic conditions.

NOTATIONS

In this paper, we follow the notation system from Kolda et al. (2009). A tensor $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}$ is defined over the real domain $\mathbb{R}$, where $d_1, d_2, \ldots, d_N$ represent the size of each dimension. Scalars, vectors, matrices, and tensors are denoted by lowercase letters (e.g., a), bold lowercase letters (e.g., $\mathbf{a} \in \mathbb{R}^M$), bold uppercase letters (e.g., $\mathbf{A} \in \mathbb{R}^{M \times N}$), and calligraphic letters (e.g., $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$), respectively. To facilitate tensor manipulation in low-rank optimization, we employ mode-n unfolding and folding operations. Mode-n unfolding transforms a tensor $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}$ into a matrix $\mathbf{X}_{(n)} \in \mathbb{R}^{d_n \times \prod_{i \neq n} d_i}$, while folding reverses this process, restoring $\mathbf{X}_{(n)}$ to its original tensor structure $\mathcal{X}$. These operations are fundamental for Schatten norm computation and truncation mechanisms in the proposed method.

PROBLEM DESCRIPTON

We use the urban road network dataset from Guangzhou, China, for problem description(OpenITS Org., 2021). The dataset is modeled as a three-dimensional tensor $\mathcal{T} \in \mathbb{R}^{214 \times 61 \times 144}$. Each tensor element $\mathcal{T}_{ijk}$ indicates the traffic speed of the i-th road segment on the j-th day within the k-th interval. Regular patterns of traffic speed are visualized through a slice along two time-related dimensions.

As shown in Figure 1, distinct temporal patterns are readily observed. On weekdays, traffic speed noticeably declines during peak hours (e.g., 7:00–9:00 and 16:30–18:30), reflecting commuting-induced congestion. In contrast, weekends (e.g., August 6, 7, and 13) and holidays (e.g., September 15–17, during the Mid-Autumn Festival) exhibit higher overall speeds, indicating reduced traffic pressure. Weekend traffic peaks are also less pronounced, with relatively stable speeds throughout the day, suggesting more dispersed travel schedules and lighter traffic loads.

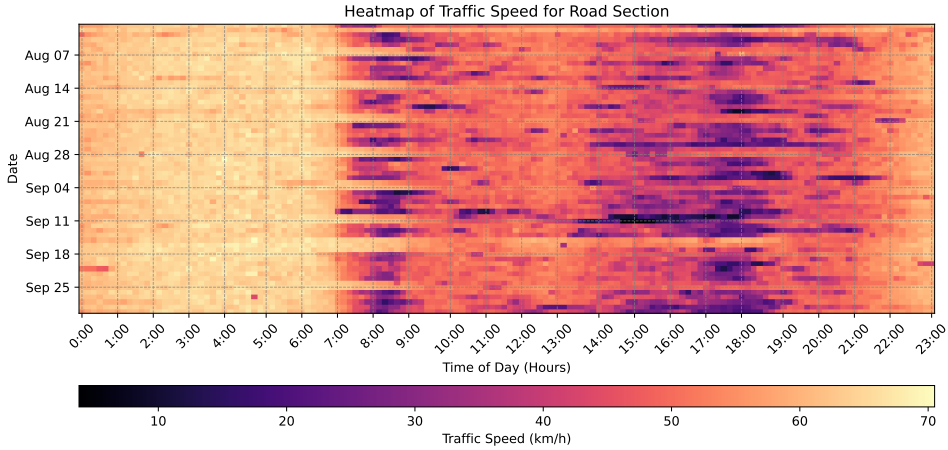


Figure 1: Time Slice Features of the G Dataset

These patterns are closely tied to the inherent spatiotemporal correlations in traffic data. Adjacent sensors often exhibit similar trends, and data from the same interval on nearby days typically align. Such local and periodic similarities are a prerequisite for filling data gaps, as they preserve spatiotemporal coherence and enable the imputed results to faithfully reflect real-world traffic dynamics.

MISSING PATTERNS

The types of missing data significantly impact traffic data restoration, influencing the complexity and applicability of imputation methods. Based on their characteristics, we categorized missing data types into three main classes: element missing (EM), fiber missing (FM), and mixed missing (MM), as depicted in Figure 2.

Element Missing

EM arises from brief sensor malfunctions or transmission errors, causing scattered data loss across road segments, days, or intervals. Due to its localized nature, EM is relatively straightforward to recover using interpolation-based methods. To simulate EM, a random tensor $R \sim U(0, 1)$, with the same dimensions as the traffic data $\mathcal{T} \in \mathbb{R}^{I \times J \times K}$, is generated. Each element $R_{ijk} \leq MR$ is then set to NaN, ensuring approximately $MR \times 100\%$ of the data is randomly omitted.

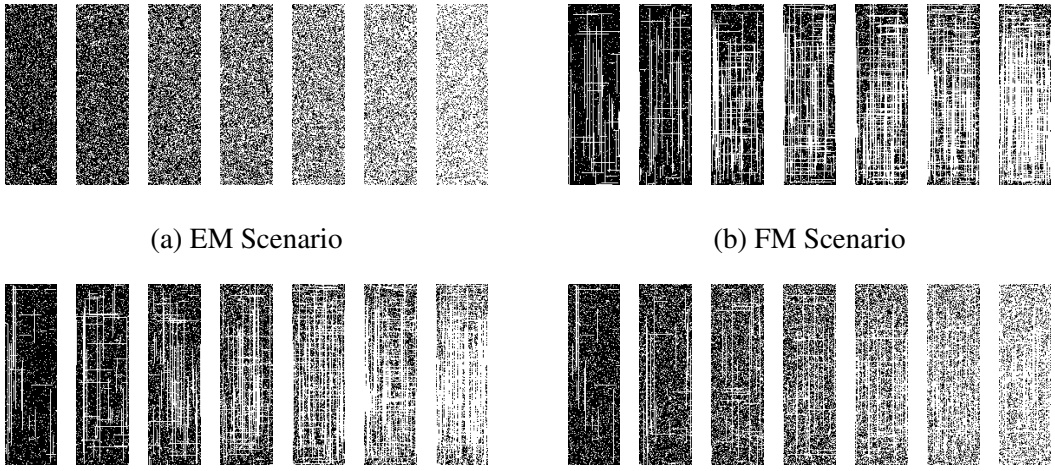
Fiber Missing

FM manifests as the loss of data for specific road segments or consecutive periods, commonly caused by prolonged sensor malfunctions or periodic system issues. It creates large-scale, sequential gaps, making recovery more difficult. FM is generated by sequentially omitting l consecutive elements along a randomly chosen dimension $d \in \{0, 1, 2\}$ (road segments, days, or intervals) of the traffic data $\mathcal{T} \in \mathbb{R}^{I \times J \times K}$, starting at a random point s_d, and repeating until the target missing rate is achieved.

Mixed Missing

MM combines EM and FM, reflecting real-world complexities such as sensor instability, transmission disruptions, and prolonged equipment failures. Its composite

nature poses the greatest challenge for data recovery, requiring methods that handle both local interpolation and large-scale restoration. To generate MM, we define the total missing rate $TMR = EMR + FMR$, where EMR and FMR represent the rates of element missing and fiber missing, respectively. By adjusting the combination of EMR and FMR, diverse missing scenarios can be easily simulated.



(a) EM Scenario

(b) FM Scenario

(c) MM (0.2EM, 0.8FM) Scenario

(d) MM (0.4EM, 0.6FM) Scenario

Figure 2: Tensor slice variations under different missing patterns with missing rates ranging from 20% to 80%

METHODOLOGY

To tackle the challenge of traffic data reconstruction under complex missing scenarios, this study proposes an Adaptive and Truncated Schatten Norm Low-Rank Tensor Completion model (LRTC-ATSN). By combining Schatten norms with a truncation mechanism, LRTC-ATSN effectively penalizes small singular values to achieve precise sparsity control. Furthermore, the ADAptive Nesterov momentum algorithm (Adan) is employed to dynamically adjust model parameters based on dataset characteristics, ensuring both efficiency and accuracy in data imputation.

Conventional LRTC

The target of LRTC is to minimize the tensor rank, formulated as $\min_{\mathcal{X}} \text{rank}(\mathcal{X})$. Since this problem is NP-hard, it is typically approximated using nuclear norm relaxation. For a tensor $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}$, the nuclear norm is defined as the sum of singular values of its mode-k unfolding $\mathbf{X}_{(k)}$, given by $\|\mathcal{X}\|_* = \sum_{k=1}^{N} \|\mathbf{X}_{(k)}\|_*$. To constrain the optimization to observed entries, an observation mask tensor $\mathcal{P}_\Omega$ is applied, retaining values in the observed set Ω while setting others to 0. The LRTC objective then minimizes the nuclear norm while ensuring that the completed tensor matches the original tensor $\mathcal{T}$ at the observed positions, expressed as:

$$\min_{\mathcal{X}} \sum_{k=1}^{N} \|\mathbf{X}_{(k)}\|_*, \quad \text{s.t.} \quad \mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{T}). \tag{1}$$

Schatten Norm

To enhance the accuracy and flexibility of LRTC, we introduce the Schatten norm as a replacement for traditional nuclear norm relaxation. For an N-th order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, the Schatten norm is defined as (Gao et al., 2020):

$$\|\mathcal{X}\|_{S_p} = \left(\sum_{k=1}^{N} \alpha_k \|\mathcal{X}_{(k)}\|_{S_p}^p \right)^{\frac{1}{p}}, \tag{2}$$

where α_k are the mode weights, representing the contribution of mode k to the tensor's overall low-rank property (Wang et al., 2022). The weights satisfy $\alpha_k \geq 0$ and $\sum_{k=1}^{N} \alpha_k = 1$. The parameter p controls the Schatten norm's sensitivity to singular values. When $p = 1$, it reduces to the convex and numerically tractable nuclear norm $\|\mathcal{X}\|_{S_1} = \sum_{i=1}^{r} \sigma_i$. For $0 < p < 1$, it becomes non-convex, effectively penalizing smaller singular values to promote greater sparsity and a more compact low-rank structure. As p approaches 0 (Figure 3), the Schatten norm closely approximates the tensor rank, improving solution accuracy at the cost of higher computational complexity.

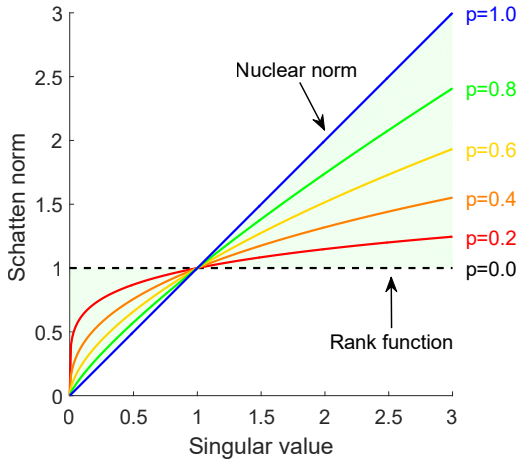


Figure 3: Curve of the Schatten Norm with Singular Values for $0 < p < 1$

Truncation Operation

Singular values quantify the importance of feature dimensions: larger singular values capture primary trends and periodic components, while smaller ones correspond to short-term, sporadic, or irregular fluctuations. To improve accuracy and reduce computational overhead, we apply a truncation operation to suppress the noise introduced by smaller singular values. This is achieved through the following formulation:

$$\|\mathcal{X}\|_{\theta,S_p} = \left(\sum_{k=1}^{N} \alpha_k \|\mathcal{X}_{(k)}\|_{\theta,S_p}^p \right)^{\frac{1}{p}} = \left(\sum_{k=1}^{N} \alpha_k \left(\sum_{i=r_k+1}^{\min\{I_k, \prod_{i \neq k} I_i\}} \sigma_i^p(\mathcal{X}_{(k)}) \right) \right)^{\frac{1}{p}}, \tag{3}$$

where r_k is the truncation threshold for each unfolded matrix, defined as:

$$r_k = \left\lceil \theta \cdot \min\{I_k, \prod_{i \neq k} I_i\} \right\rceil, \quad \forall k \in \{1, 2, \ldots, N\}, \tag{4}$$

with $\theta \in [0, 1]$ being the global truncation rate, and $\lceil \cdot \rceil$ indicating the ceiling function. The singular values σ_i are sorted in non-increasing order.

This operation preserves the largest r_k singular values while penalizing the remaining $\left(\min\{I_k, \prod_{i \neq k} I_i\} - r_k\right)$ smaller ones. Integrating the truncated Schatten norm into the LRTC framework, the optimization problem becomes:

$$\min_{\mathcal{X}} \|\mathcal{X}\|^p_{\theta,S_p} = \min_{\mathcal{X}} \sum_{k=1}^{N} \alpha_k \|\mathcal{X}_{(k)}\|^p_{\theta,S_p}, \quad \text{s.t.} \quad \mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{T}). \tag{5}$$

Adaptive and Smoothing Update Mechanism

Choosing optimal parameters p, truncation rate θ, and mode weights α_k is critical yet tricky, as they vary based on the dataset, missing patterns, and iteration stage. To balance the accuracy and efficiency of imputation, we employ the Adan algorithm for adaptive updates of p and θ, while a regularization term and weighted moving average method are applied to smooth the updates of α_k (Zhao et al., 2024).

Adaptive update of p ***and*** θ

At each iteration t, we compute the error gradient $g^{(t)} = E^{(t)} - E^{(t-1)}$ and update the first- and second-order momenta, $m(t)$ and $v(t)$, as follows:

$$m^{(t)} = \beta_1 m^{(t-1)} + (1 - \beta_1) g^{(t)}, \quad v^{(t)} = \beta_2 v^{(t-1)} + (1 - \beta_2) \left(g^{(t)}\right)^2, \tag{6}$$

where β_1 and β_2 are the decay coefficients for the two momenta, typically set to $\beta_1 = 0.9$ and $\beta_2 = 0.999$ (Xie et al., 2024). To mitigate initial bias, we correct the momentum terms to better reflect the true gradient values:

$$\hat{m}^{(t)} = \frac{m^{(t)}}{1 - \beta_1^t}, \quad \hat{v}^{(t)} = \frac{v^{(t)}}{1 - \beta_2^t}. \tag{7}$$

Using the corrected momentum terms, we can adaptively update p and θ:

$$p_{t+1} = \text{clip}\left(p_t - \eta \frac{\hat{m}^{(t)}}{\sqrt{\hat{v}^{(t)}} + \epsilon},\ 0.1,\ 1\right), \quad \theta_{t+1} = \text{clip}\left(\theta_t + \eta \frac{\hat{m}^{(t)}}{\sqrt{\hat{v}^{(t)}} + \epsilon},\ 0,\ 1\right), \tag{8}$$

here, the learning rate η controls the update step size, ϵ is a small constant to avoid division by 0, and function $\text{clip}(\cdot)$ constrains p and θ within valid ranges.

Smoothing update of α_k

The update of the mode weights α_k is performed by calculating the Frobenius norm for each mode, defined as $s_k = \|\mathcal{M}_k\|_F$, and introducing a regularization term to specify normalized candidate weights:

$$\tilde{\alpha}_k = \frac{s_k}{\sum_{j=1}^{N} s_j} + \lambda \left(\frac{1}{N} - \frac{s_k}{\sum_{j=1}^{N} s_j}\right), \tag{9}$$

where λ is the regularization coefficient that controls shrinkage degree towards the average, preventing any mode weight from dominating excessively.

We use a weighted moving average to smooth weight updates by combining historical and current information, reducing abrupt changes. The method is given by:

$$\alpha_k^{(t+1)} = (1-\gamma)\alpha_k^{(t)} + \gamma\tilde{\alpha}_k, \tag{10}$$

followed by normalization, $\alpha_k^{(t+1)} = \alpha_k^{(t+1)} / \sum_{j=1}^{N} \alpha_j^{(t+1)}$, where γ is the learning rate for weight updates.

SOLUTION ALGORITHM

To efficiently solve the truncated tensor Schatten norm minimization problem, we adopt an algorithm combining the Alternating Direction Method of Multipliers (ADMM) (Zeng et al., 2024) with Truncated Generalized Soft Thresholding (TGST) (Goulart et al., 2017), referred to as ADMM-TGST. This approach breaks the complex optimization problem into simpler subproblems and iteratively approximates the optimal solution. The solution process is depicted in Figure 4.

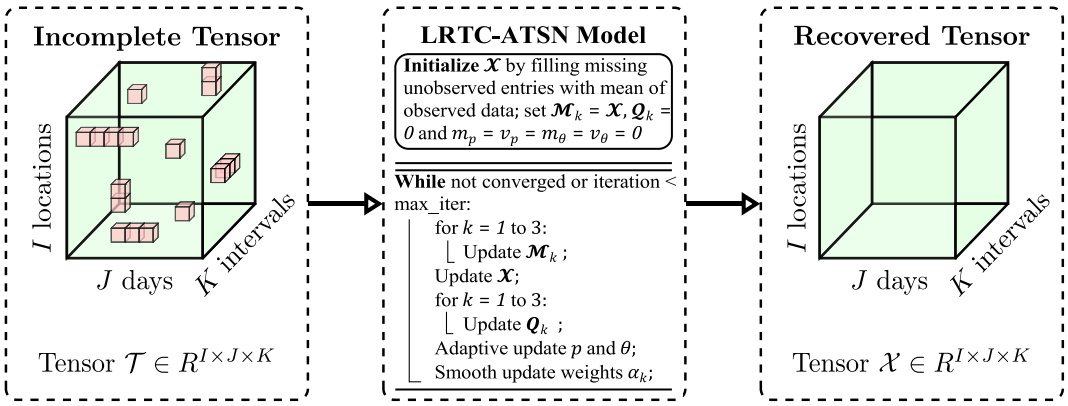


Figure 4: Solution Process of the LRTC-ATSN Model

To facilitate the solution process, auxiliary variables $\{\mathcal{M}_k\}_{k=1}^N$ are employed to reformulate the original optimization problem (Equation (5)) as follows:

$$\min_{\mathcal{X},\{\mathcal{M}_k\}} \quad \sum_{k=1}^{N} \alpha_k \left\|\mathcal{M}_k\right\|_{\theta,S_p}^p, \quad \text{s.t.} \quad \begin{cases} \mathcal{M}_k = \mathcal{X}, & \forall k = 1,2,\ldots,N, \\ \mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{T}). \end{cases} \tag{11}$$

Building on this, we introduce Lagrange multipliers $\{\mathcal{Q}_k\}_{k=1}^N$ and construct the augmented Lagrangian function:

$$\mathcal{L}(\mathcal{X},\{\mathcal{M}_k\},\{\mathcal{Q}_k\}) = \sum_{k=1}^{N}\left[\alpha_k\|\mathcal{M}_k\|_{\theta,S_p}^p + \langle\mathcal{Q}_k, \mathcal{X}-\mathcal{M}_k\rangle + \frac{\beta}{2}\left\|\mathcal{X}-\mathcal{M}_k\right\|_F^2\right], \tag{12}$$

here, $\beta > 0$ is the penalty parameter for the Lagrangian multiplier enforcing equality constraints. The inner product $\langle\mathcal{A},\mathcal{B}\rangle$ defined as $\sum_{i_1,i_2,\ldots,i_N} a_{i_1,i_2,\ldots,i_N} b_{i_1,i_2,\ldots,i_N}$.

Subsequently, we apply ADMM to alternately optimize the variables $\mathcal{X}$, $\mathcal{M}_k$, and $\mathcal{Q}_k$. The detailed optimization steps are outlined below:

Updating $\mathcal{M}_k$

To solve for $\mathcal{M}_k$, ADMM alternates optimization by fixing $\mathcal{X}^{(t)}$ and $\mathcal{Q}_k^{(t)}$ during the t-th iteration. The objective minimizes the Schatten norm regularization term and the penalty on $\mathcal{M}_k$, expressed as:

$$\mathcal{M}_k^{(t+1)} = \arg\min_{\mathcal{M}_k} \ \alpha_k \|\mathcal{M}_k\|_{\theta,S_p}^p + \left\langle \mathcal{Q}_k^{(t)}, \mathcal{X}^{(t)} - \mathcal{M}_k \right\rangle + \frac{\beta}{2} \left\| \mathcal{X}^{(t)} - \mathcal{M}_k \right\|_F^2. \quad (13)$$

By expanding the linear term and rearranging, this can be simplified to:

$$\mathcal{M}_k^{(t+1)} = \arg\min_{\mathcal{M}_k} \ \alpha_k \|\mathcal{M}_k\|_{\theta,S_p}^p + \frac{\beta}{2} \left\| \mathcal{M}_k - \left(\mathcal{X}^{(t)} + \frac{1}{\beta} \mathcal{Q}_k^{(t)} \right) \right\|_F^2. \quad (14)$$

Unfold the tensor $\mathcal{M}_k$ and $\mathcal{X}^{(t)} + \frac{1}{\beta}\mathcal{Q}_k^{(t)}$ along the k-th mode into matrices $\mathbf{M}_{(k)}$ and $\mathbf{Y}_k$, respectively. Perform singular value decomposition (SVD) on $\mathbf{Y}_k = \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^\top$, where $\mathbf{\Sigma}_k = \mathrm{diag}(\sigma_1, \ldots, \sigma_R)$.

To enforce low-rank properties, a truncation threshold $r_k = \lceil \theta \cdot R \rceil$ selects r_k dominant singular values to retain. Singular values beyond r_k are adjusted through Generalized Soft Thresholding (GST):

$$\delta_i = \mathrm{GST}\left(\sigma_i, \tau_{k,i}, p\right), \quad \tau_{k,i} = \frac{\alpha_k w_{k,i}}{\beta}, \quad w_{k,i} = \begin{cases} 0, & i \le r_k, \\ 1, & i > r_k. \end{cases} \quad (15)$$

Reconstruct $\mathbf{\Delta}_k = \mathrm{diag}(\delta_1, \ldots, \delta_R)$, update $\mathbf{M}_{(k)}^{(t+1)} = \mathbf{U}_k \mathbf{\Delta}_k \mathbf{V}_k^\top$, and fold it back into tensor form $\mathcal{M}_k^{(t+1)}$ to complete the update step.

Updating $\mathcal{X}$

To determine the optimal $\mathcal{X}$ while keeping $\mathcal{M}_k^{(t+1)}$ and $\mathcal{Q}_k^{(t)}$ fixed, we formulate the following minimization problem to solve for $\mathcal{X}^{(t+1)}$:

$$\mathcal{X}^{(t+1)} = \arg\min_{\mathcal{X}} \sum_{k=1}^{N} \left[\left\langle \mathcal{Q}_k^{(t)}, \mathcal{X} - \mathcal{M}_k^{(t+1)} \right\rangle + \frac{\beta}{2} \left\| \mathcal{X} - \mathcal{M}_k^{(t+1)} \right\|_F^2 \right]. \quad (16)$$

The constraint is satisfied by the analytical solution given below:

$$\mathcal{X}^{(t+1)} = \mathcal{P}_\Omega(\mathcal{T}) + \mathcal{P}_{\bar{\Omega}} \left(\frac{1}{N} \sum_{k=1}^{N} \left(\mathcal{M}_k^{(t+1)} - \frac{1}{\beta} \mathcal{Q}_k^{(t)} \right) \right). \quad (17)$$

where $\mathcal{P}_{\bar{\Omega}} = \mathbf{I} - \mathcal{P}_\Omega$ is the projection operator onto the unobserved entries, and $\mathbf{I}$ is the identity operator. This step updates $\mathcal{X}$ by integrating information from all modes while ensuring the observed data remains unchanged.

Updating $\mathcal{Q}_k$

To approximate the ideal solution satisfying the equality constraint $\mathcal{X} = \mathcal{M}_k$, we update the Lagrange multipliers $\mathcal{Q}_k$ to ensure consistency between $\mathcal{X}$ and $\mathcal{M}_k$:

$$\mathcal{Q}_k^{(t+1)} = \mathcal{Q}_k^{(t)} + \beta \left(\mathcal{X}^{(t+1)} - \mathcal{M}_k^{(t+1)} \right). \quad (18)$$

Each $\mathcal{Q}_k$ is updated for every mode k, gradually enhancing the similarity between $\mathcal{X}$ and $\mathcal{M}_k$, ultimately stabilizing $\mathcal{X}$ in a state that satisfies the constraint.

Iteration and Convergence

Continue updating the auxiliary variables $\mathcal{M}_k$, the primary variable $\mathcal{X}$, and the Lagrange multipliers $\mathcal{Q}_k$ until the following condition is met:

$$\frac{\|\mathcal{X}^{(t+1)} - \mathcal{X}^{(t)}\|_F}{\|\mathcal{X}^{(t)}\|_F} < \epsilon, \tag{19}$$

where $\epsilon > 0$ is the predefined accuracy threshold which determines when the algorithm has converged to a satisfactory solution.

EXPERIMENTS

Datasets

To validate the LRTC-ATSN model's performance in real-world scenarios, two datasets from actual traffic systems were utilized. Although their dimensions differ, both share an identical tensor structure: location × day × interval.

a) Guangzhou Urban Road Network Speed Dataset (G): Traffic speed data recorded from August 1 to September 30, 2016, at 10-minute intervals (144 time windows/day), covering 214 urban expressways and main roads with a 1.29% missing rate. The tensor dimensions are $214 \times 61 \times 144$.

b) Seattle Freeway Traffic Speed Dataset (S) (Cui et al., 2019): Traffic speed data recorded from January 1 to January 28, 2015, at 5-minute intervals (288 time windows/day), collected from 323 loop detectors on major freeways in Seattle. The tensor dimensions are $323 \times 28 \times 288$.

Baseline Models

We compared the proposed method with five representative baseline models: HaLRTC (Liu et al., 2012), LRTC-TNN (X. Chen, J. Yang, et al., 2020), LRTC-TSpN (Nie et al., 2022), BGCP (X. Chen, He, et al., 2019), and BPMF (Vander Aa et al., 2017). Among these, HALRTC, LRTC-TNN, and LRTC-TSpN are LRTC models that achieve low-rank approximations through nuclear norm, truncated nuclear norm, and truncated Schatten norm, respectively. BGCP and BPMF are grounded in Bayesian inference, with the former using Gaussian CP decomposition for time-dependent data recovery and the latter leveraging probabilistic matrix factorization to effectively describe latent structures, making it suitable for large-scale matrix completion.

Evaluation Metrics

Data imputation accuracy is assessed using MAPE and RMSE. MAPE measures the average relative error, while RMSE reflects the root of the average squared error. Lower MAPE and RMSE values indicate better model performance.

Experimental Results

Convergence analysis

Computational complexity remains a major challenge in non-convex optimization problem. To test the efficiency of the LRTC-ATSN solving algorithm, we conducted a convergence analysis on datasets G and S under three complex missing patterns (EM-FM combinations: 0.1/0.3, 0.2/0.2, 0.3/0.1), as shown in Figure 5.

The results show that MAPE and RMSE drop rapidly within the initial 5 iterations and stabilize at an acceptable level by the 30 iterations, emphasizing the effectiveness and applicability of the ADMM-TGST optimization scheme across scenarios.

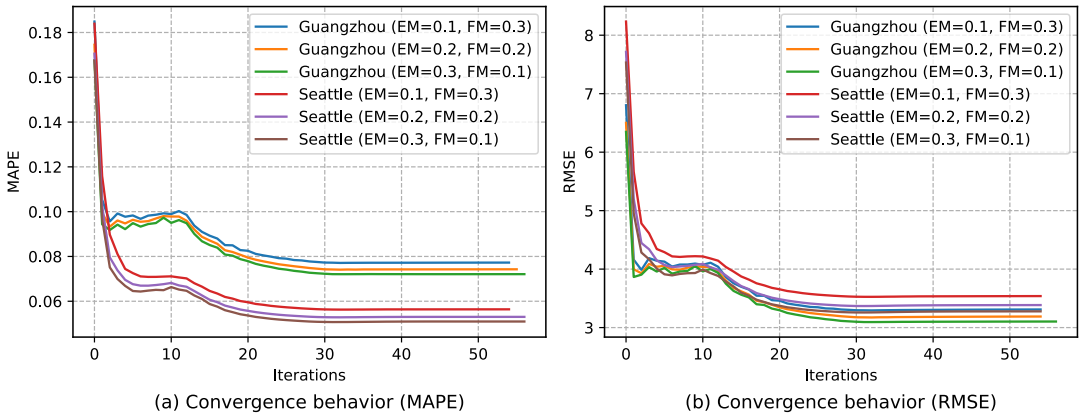


Figure 5: Convergence Performance of the Model under Different Scenarios

Performance comparison

Table 1 presents the performance of different models under various missing patterns for the G and S datasets. The results include combinations of EM-FM at rates of 0.1, 0.3, 0.5, 0.7, and 0.9. The evaluation metrics used are MAPE / RMSE.

Table 1: Performance comparison of different models on various datasets

Datasets	Missing Pattern	TMR	BPMF	BGCP	HaLRTC	LRTC-TNN	LRTC-TSpN	LRTC-ATSN
G	EM 0.1 FM 0.1	19.60%	0.106/4.412	0.104/4.321	0.088/3.561	0.073/3.160	0.073/3.132	**0.072/3.097**
	EM 0.3 FM 0.3	48.62%	0.108/4.491	0.104/4.318	0.098/3.929	0.082/3.530	0.081/3.423	**0.078/3.360**
	EM 0.5 FM 0.5	69.75%	0.112/4.665	0.103/4.316	0.109/4.301	0.089/3.849	0.087/3.689	**0.087/3.676**
	EM 0.7 FM 0.7	85.00%	0.119/4.951	0.104/4.320	0.123/4.800	0.098/4.195	0.093/**3.943**	**0.093**/3.947
	EM 0.9 FM 0.9	95.85%	0.139/5.675	0.107/4.473	0.819/33.808	0.111/4.627	0.105/4.409	**0.105/4.355**
S	EM 0.1 FM 0.1	18.52%	0.098/5.526	0.100/5.613	0.071/3.952	0.052/3.329	0.052/3.318	**0.050/3.233**
	EM 0.3 FM 0.3	47.91%	0.099/5.616	0.101/5.615	0.082/4.440	0.060/3.761	0.058/3.583	**0.057/3.582**
	EM 0.5 FM 0.5	69.37%	0.103/5.818	0.099/5.555	0.098/5.078	0.072/4.359	**0.065/3.966**	0.066/4.028
	EM 0.7 FM 0.7	84.80%	0.114/6.327	0.101/5.640	0.121/5.959	0.090/5.274	0.077/**4.531**	**0.076**/4.590
	EM 0.9 FM 0.9	95.78%	0.169/8.529	0.105/5.900	0.200/9.356	0.360/23.742	0.113/6.135	**0.101/5.758**

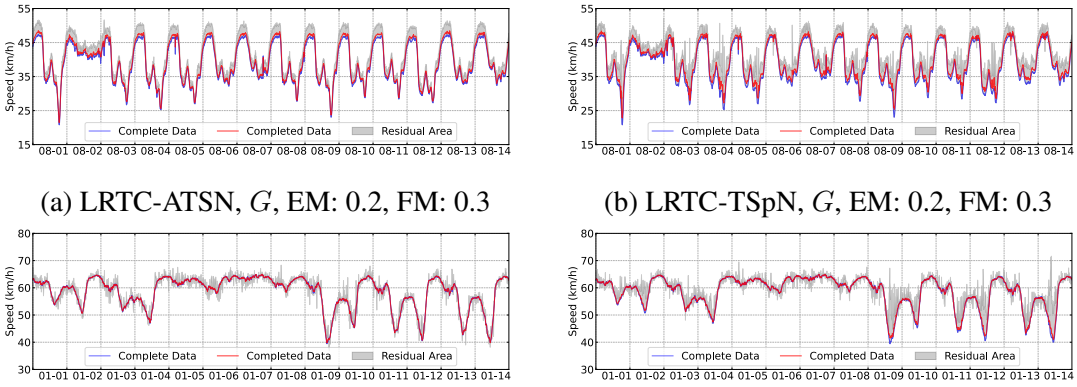
Best results are marked in bold font.

The results indicate that LRTC-ATSN consistently outperforms other models, particularly at higher missing rates (e.g., 85.00% and 95.85%). Its superior imputation precision is evidenced by its steadily lower MAPE and RMSE values.

Imputation Visualization

The LRTC-ATSN and LRTC-TSpN models share some similarities in data imputation methodologies, both employing truncation and non-convex optimization. Nevertheless, their implementations differ. To evaluate the stability and robustness of the adaptive mechanism in LRTC-ATSN under complex missing patterns, comparative experiments were conducted. For the G dataset, an EM of 0.2 and an FM of 0.3 (scenario a) produced approximately 41.28% MM data. For the S dataset, an EM of 0.5 and an FM of 0.6 (scenario b) resulted in 72.21% MM data.

Figure 6 compares actual and imputed values, with the residual area magnified for clarity. In scenario a, LRTC-ATSN outperforms LRTC-TSpN, reducing the residual



(a) LRTC-ATSN, G, EM: 0.2, FM: 0.3

(b) LRTC-TSpN, G, EM: 0.2, FM: 0.3

(c) LRTC-ATSN, S, EM: 0.5, FM: 0.6

(d) LRTC-TSpN, S, EM: 0.5, FM: 0.6

Figure 6: Imputation Performance of LRTC-ATSN vs. LRTC-TSpN

area by 28.3% (1463.12 vs. 2040.98). In scenario b, with higher data variability and missing rates, LRTC-ATSN shows a 36.1% improvement (320.33 vs. 501.29). This greater capability is due to its dynamic parameter and weight adjustments, enabling LRTC-ATSN to maintain realistic fit recovery curves even under high missing rates.

CONCLUSION

In this study, we present the Adaptive and Truncated Schatten Norm Low-Rank Tensor Completion (LRTC-ATSN) model, a novel approach for traffic data completion under complex missing patterns. By integrating dynamic parameter adjustment and adaptive truncation, the proposed model significantly improves imputation accuracy and efficiency. To thoroughly evaluate its performance, we developed a framework for generating missing patterns that closely simulate realistic traffic scenarios. Experimental results on real-world traffic datasets reveal that LRTC-ATSN consistently outperforms existing models. At an extreme missing rate of 95.78%, it achieves a 10.6% reduction in MAPE and a 6.1% reduction in RMSE compared to the best-performing baseline. These results highlight the potential of the LRTC-ATSN model to enhance data reliability in ITS. To further advance the model's effectiveness, future work will focus on optimizing the adaptive mechanism to reduce computational costs and extending the model's application to other domains requiring robust data imputation.

ACKNOWLEDGEMENTS

This work was supported by the Key Research Platforms and Projects of the Department of Education of Guangdong Province (Nos. 2022ZDZX1021 & 2024ZDZX1033), and the Research Project of the Bureau of Education of Guangzhou Municipality (No. 2024312023).

REFERENCES

Chen, Xinyu, Zhanhong Cheng, et al. (2024). "Laplacian convolutional representation for traffic time series imputation". In: *IEEE Transactions on Knowledge and Data Engineering*.

Chen, Xinyu, Zhaocheng He, et al. (2019). "A Bayesian tensor decomposition approach for spatiotemporal traffic data imputation". In: *Transportation research part C: emerging technologies* 98, pp. 73–84.

Chen, Xinyu, Jinming Yang, et al. (2020). "A nonconvex low-rank tensor completion model for spatiotemporal traffic data imputation". In: *Transportation Research Part C: Emerging Technologies* 117, p. 102673.

Chen, Yong et al. (2022). "A novel reinforced dynamic graph convolutional network model with data imputation for network-wide traffic flow prediction". In: *Transportation Research Part C: Emerging Technologies* 143, p. 103820.

Cui, Zhiyong et al. (2019). "Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting". In: *IEEE Transactions on Intelligent Transportation Systems* 21.11, pp. 4883–4894.

Gao, Shangqi et al. (2020). "Robust Schatten-p norm based approach for tensor completion". In: *Journal of Scientific Computing* 82, pp. 1–23.

Goulart, JH de M et al. (2017). "Traffic data imputation via tensor completion based on soft thresholding of Tucker core". In: *Transportation Research Part C: Emerging Technologies* 85, pp. 348–362.

Kolda, Tamara G et al. (2009). "Tensor decompositions and applications". In: *SIAM review* 51.3, pp. 455–500.

Liu, Ji et al. (2012). "Tensor completion for estimating missing values in visual data". In: *IEEE transactions on pattern analysis and machine intelligence* 35.1, pp. 208–220.

Nie, Tong et al. (2022). "Truncated tensor Schatten p-norm based approach for spatiotemporal traffic data imputation with complicated missing patterns". In: *Transportation research part C: emerging technologies* 141, p. 103737.

OpenITS Org. (2021). *OpenData V12.0—Large-scale Traffic Speed Data Set*. [Online]. Available: `https://www.openits.cn/openData2/792.jhtml`. (Visited on 11/05/2024).

Vander Aa, Tom et al. (2017). "Distributed Bayesian probabilistic matrix factorization". In: *Procedia Computer Science* 108, pp. 1030–1039.

Wang, Shaofan et al. (2022). "Spatiotemporal traffic data imputation via tensorial weighted Schatten-p norm minimization". In: *IET Intelligent Transport Systems* 16.7, pp. 926–939.

Xie, Xingyu et al. (2024). "Adan: Adaptive nesterov momentum algorithm for faster optimizing deep models". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Yang, Lei et al. (2024). "A Brief Review on Missing Traffic Data Imputation Methods for Intelligent Transportation Systems". In: *2024 7th International Symposium on Autonomous Systems (ISAS)*. IEEE, pp. 1–6.

Zeng, Zeyu et al. (2024). "Low-rank tensor and hybrid smoothness regularization-based approach for traffic data imputation with multimodal missing". In: *IEEE Transactions on Intelligent Transportation Systems*.

Zhao, Meng et al. (2024). "Change Detection in Partially Observed Large-Scale Traffic Network Data". In: *IEEE Transactions on Intelligent Transportation Systems*.