Fourier-Enhanced Adaptive Manifold Latent Feature Analysis for Spatiotemporal Signal Recovery

Yuting Ding[1], Jianyong Zheng[2]✉, Fei Mei[3], and Ang Gao[2]

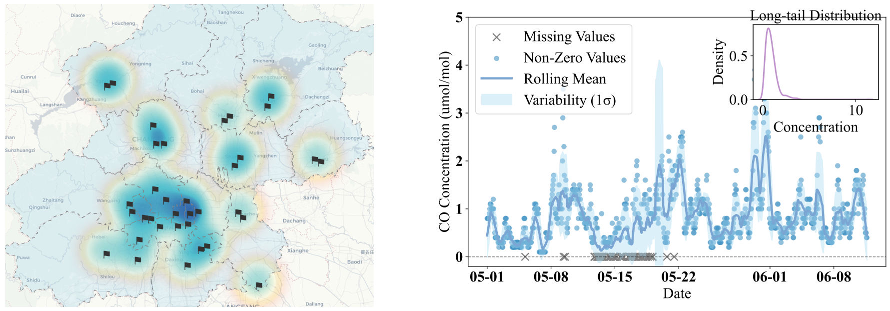
[1] School of Cyber Science and Engineering, Southeast University, Nanjing, China
ading_ding@163.com
[2] School of Electrical Engineering, Southeast University, Nanjing, China
jyzheng209@163.com, 220202894@seu.edu.cn
[3] College of Electrical and Power Engineering, Hohai University, Nanjing, China
meifei@hhu.edu.cn

Abstract. Wireless Sensor Networks (WSNs) face critical data incompleteness challenges driven by hardware failures and energy constraints, which severely undermine environmental monitoring reliability. Although are frequently employed, Low-Rank Matrix Approximation (LRMA) methods often overlook nonlinear temporal dynamics and fail to discriminate structured noise from actual anomalies. This paper introduces the Adaptive Latent Feature Analysis with Fourier Embedding (ALFA-FE) framework, featuring two principal contributions: (1) dynamic Fourier embeddings that incorporate manifold-based frequency-domain regularization to flexibly capture multi-scale temporal patterns, and (2) a robust optimization scheme unifying Huber-norm loss with anomaly-sensitive constraints. Comprehensive evaluations across four real-world datasets reveal that ALFA-FE significantly outperforms seven state-of-the-art models in both reconstruction accuracy and robustness. By effectively balancing precise signal recovery with anomaly retention, ALFA-FE demonstrates strong potential for advancing environmental sensing reliability in resource-limited IoT deployments.

Keywords: Wireless sensor networks(WSNs)· Low-rank matrix approximation(LRMA)· Manifold Regularization· Spatio-temporal correlation.

1 Introduction

WSNs consist of a deluge of distributed sensor nodes that communicate wirelessly to monitor and collect environmental data [28]. Each sensor node is capable of performing preliminary data processing, such as filtering, aggregation, and fusion, to reduce redundant data and enhance data transmission efficiency [5]. With the continuous evolution of WSNs, adoption has expanded across diverse domains, including environmental monitoring, smart cities [30, 4, 6], healthcare management [9, 34], and industrial automation [19, 32]. WSNs enable real-time, high-resolution data acquisition, playing a crucial role in enhancing decision-making and optimizing resource utilization across various fields. The significant



(a) PM2.5 (b) CO

Fig. 1: The data distribution of real Beijing air quality.

advantages of WSNs have established it as a key driver of efficiency and innovation in multiple domains.

WSNs often suffer from data loss due to sensor failures, communication disruptions, and changes in network topology [10]. Ensuring reliable data collection becomes particularly challenging when monitoring dynamic environments, where measurements must be continuously captured and transmitted [33, 29]. Moreover, the data collected by WSNs often exhibit complex spatiotemporal dependencies, influenced by environmental cycles or periodic industrial operations, further complicating the recovery process [18]. To tackle this issue effectively, missing data appears as an incomplete matrix, with rows corresponding to sensor nodes and columns representing time slots. Given the inherent correlations within WSN data, including nonlinear dependencies across both spatial and temporal dimensions, the resulting matrix often exhibits low-rank structure. Consequently, recovering missing values can be formulated as LRMA problem, where the missing entries are inferred based on the relationships embedded in the observed data.

LMRA establishes a mathematical framework for matrix decomposition, employing rank constraints to construct low-dimensional manifold approximations of high-dimensional data [37]. Derived from this theory, low-rank constrained optimization models enable effective missing data recovery in wireless sensor networks by leveraging subspace structures of partial observations [21, 37, 26]. However, current implementations ignore the intrinsic heterogeneity of anomalies across multisource sensors [12], characterized by divergent spatiotemporal signatures across sensor clusters. To be more intuitive, Fig. 1 shows a real example to demonstrate this issue.

Example 1: Air quality monitoring data from Beijing WSNs (5.1-6.8, 2014) demonstrate sensor challenges. Fig. 1a reveals intensified urban gradients that indicate persistent high-concentration variability, while the suburban site remains comparatively stable. Fig. 1b indicates a multimodal CO distribution ($\mu = 1.8$, $\sigma = 0.5\,\mathrm{mg/m^3}$) with 9.4% of samples above $2.7\,\mathrm{mg/m^3}$. Three transient spikes exceeding $3.5\,\mathrm{mg/m^3}$ appear around May 25–28, peaking at $3.8\,\mathrm{mg/m^3}$, contrasting the typical diurnal cycle of 1.1–$2.3\,\mathrm{mg/m^3}$. Spatial-temporal coupling between PM2.5 volatility and CO anomalies suggests distinct noise-outlier interactions.

Example 1 substantiates the challenges of strongly coupled outliers and pervasive multimodal noise within data gathered by WSNs. While conventional approaches predominantly resort to static regularization schemes, such methods prove inadequate as predetermined weighting parameters and rigid processing frameworks fail to accommodate the intricate, nonlinear dynamics intrinsic to noise–outlier coupling phenomena. A case in point is PM2.5, which exhibit a rapid increase and inconsistency with nearby sensor readings. Needless to say, such an outlier cannot uniformly treated as random noise. Alternatively, can we build adaptive LRMA architectures that improve recovery accuracy by achieving noise suppression and anomaly retention?

In response, this paper proposes an adaptive latent feature analysis (ALFA) framework with Fourier embedding (FE) for spatiotemporal signal recovery in WSN data, termed ALFA-FE, balancing noise resistance and anomaly retention. Its main idea is twofold: 1) temporal-focused Fourier integration enables the latent space to dynamically adapt to non-stationary patterns through manifold-basedfrequency-domain regularization. 2) Aggregating the huber-norm with gradient descent empowers the model to discern legitimate anomalies from coupled noise-artifacts. With such a design, it possesses the merits of both robustly capturing temporal dynamics and effectively distinguishing noise from genuine anomalies.

The primary contributions of this work are summarized as follows:

- An ALFA-FE model is constructed by integrating Fourier feature embedding with the Huber norm, facilitating dynamic frequency regulation and robust anomaly handling in spatiotemporal WSN data.
- Theoretical guarantees, including convergence analysis and generalization bounds, are provided for the ALFA-FE model.
- An efficient optimization algorithm is designed by synergizing gradient-based updates with adaptive noise suppression, ensuring scalability and real-time adaptability for large-scale WSN deployments.

The computational complexity analysis of ALFA-FE and some experimental detail are submiteed as supplemental material. ALFA-FE code and supplemental material have been published on https://github.com/adingyuting/signal-recovery

2 Related Work

The LRMA model formulates WSN data recovery as a low-rank matrix completion problem, leveraging its intrinsic low-rank structure to represent data with a limited set of latent factors. By harnessing spatial and temporal dependencies from observed data points, LRMA optimizes an objective function to infer missing values, ensuring a globally coherent reconstruction. In recent years, LRMA has made significant strides in WSN data recovery, integrating various regularization techniques to enhance reconstruction accuracy, including graph regularization, nuclear norm minimization, sparsity constraints, weighted L1-L2 regularization, temporal smoothness enforcement, total variation regularization, and Bayesian prior modeling.

The exploitation of spatiotemporal features constitutes the cornerstone of LRMA-based data recovery in WSNs. Prevailing methods predominantly resort to graph-based regularization to encode such dependencies—ranging from graph Laplacian matrices for topology-aware smoothing [27, 20] to hybrid graph constraints capturing global-local signal structures [8]. Despite their merits, these approaches intrinsically presume linear or quasi-linear temporal dependencies, often failing to accommodate the nonlinear dynamics pervasive in real-world WSN signals [25, 20, 7].

Regarding robustness enhancement, the community has witnessed a dichotomy between L2-norm-oriented accuracy optimization [44, 14] and L1-norm-driven outlier resistance [13, 36]. While L2-based models achieve superior performance under Gaussian noise, their susceptibility to anomalies has been widely documented [41, 38]. Conversely, L1-regularized variants, though more robust, tend to undermine recovery fidelity due to over-conservative noise suppression [15, 2, 39, 43]. Recent attempts to reconcile this dilemma through static L1/L2 hybrid regularization [45, 16] merely mitigate rather than resolve the bias-variance tradeoff, as their fixed weighting parameters cannot adapt to skewed noise distributions. This paper introduces two innovations: Nonlinear Temporal Regularization and Adaptive Robustness via Huber Regularization. The first embeds coupled Fourier features to capture nonlinear spatiotemporal dependencies, overcoming graph-based methods' linearity assumption. The second replaces rigid L1/L2 composites with Huber loss for dynamic outlier response, synergizing with nonlinear representation learning to form a theoretically grounded framework that achieves an optimal equilibrium unattainable by prior arts.

3 The proposed ALFA-FE model

The proposed ALFA-FE framework advances LFA models through dual-criteria optimization. It exploits spatiotemporal manifold regularization term that captures nonlinear time-dependent patterns and employs Huber-norm for noise suppression and anomaly preservation, offering adaptive responses via self-adjusted thresholds. The next is to introduce the proposed ALFA-FE model.

3.1 Exploiting Spatiotemporal Feature

The spatiotemporal features in WSNs are characterized by spatial proximity and temporal dynamics. Spatially, sensor measurements attenuate with increasing node distance. Temporally, sensor signals show stable frequency and diverse dynamics across different time scales.

Spatial proximity To capture spatial proximity, we construct undirected weighted sensor graph $G = (\mathcal{V}, \mathcal{E}, A)$, where $\mathcal{V}$ represents the set of N sensor nodes, $\mathcal{E}$ denotes the set of edges representing pairwise sensor connections, and

$A \in \mathbb{R}^{N \times N}$ is the weighted adjacency matrix. Each entry a_{ij} in A quantifies the relationship between sensor node i and node j. Given the sensor locations, the pairwise distance matrix $D \in \mathbb{R}^{N \times N}$ is computed based on Euclidean distances:

$$d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|_2, \tag{1}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ represent the coordinate vectors of the respective sensors.

Using the distance matrix D, the adjacency matrix A is constructed by selecting k nearest neighbors for each node. The edge weights are determined by a Gaussian kernel function:

$$a_{ij} = \begin{cases} \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right), & \text{if } j \in \mathcal{N}_k(i) \text{ or } i \in \mathcal{N}_k(j) \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where σ is a scaling parameter controlling the decay of similarity with distance.

The spatial smoothness is characterized by the graph Laplacian matrix L, defined as:

$$L = D - A, \tag{3}$$

where D is the degree matrix, a diagonal matrix whose elements are given by:

$$D_{ii} = \sum_j a_{ij}. \tag{4}$$

Thus, L is expressed as:

$$L = \begin{bmatrix} \sum a_{1j} & -a_{12} & \cdots & -a_{1N} \\ -a_{21} & \sum a_{2j} & \cdots & -a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{N1} & -a_{N2} & \cdots & \sum a_{Nj} \end{bmatrix}_{N \times N}, \tag{5}$$

where $\sum a_{ij}$ denotes the sum of weights in the i-th row of A. The Laplacian matrix encodes the structural properties of the sensor network and enables spatial proximity modeling.

To normalize the influence of varying node degrees, we further define the normalized graph Laplacian as:

$$L_{\text{sym}} = I - D^{-1/2} A D^{-1/2}, \tag{6}$$

and the random walk Laplacian as:

$$L_{rw} = I - D^{-1} A. \tag{7}$$

Normalized Laplacians ensure scale-invariance and enhance spectral analysis of the sensor network. By constructing the graph Laplacian, ALFA-FE effectively captures spatial dependencies among sensor nodes, enabling robust graph-based signal processing techniques for spatiotemporal modeling in wireless sensor networks.

Temporal dynamic WSNs measurements, represented as rows in $\mathbf{W} \in \mathbb{R}^{N\times T}$, exhibit structured temporal dynamics, often containing both long-term trends and high-frequency variations. Fixed-scale representations may fail to capture this variability, leading to either the loss of fine details or excessive sensitivity to noise. To fix this issue, this paper employs Concatenated Fourier Features (CFF), which embeds time into a multi-scale sinusoidal basis, ensuring a robust representation of diverse temporal patterns.

Given a sequence of time indices $\boldsymbol{\tau} = [\tau_1, \tau_2, \dots, \tau_T]^T \in \mathbb{R}^T$, we define the Fourier feature expansions:

$$\gamma(\boldsymbol{\tau}) = \Big[\sin(2\pi\mathbf{B}_1\boldsymbol{\tau}),\, \cos(2\pi\mathbf{B}_1\boldsymbol{\tau}),\, \dots,\, \sin(2\pi\mathbf{B}_S\boldsymbol{\tau}),\, \cos(2\pi\mathbf{B}_S\boldsymbol{\tau})\Big]^T \tag{8}$$

where $\mathbf{B}_s \in \mathbb{R}^{d/2\times 1}$ frequency scaling matrices are sampled from a Gaussian distribution, i.e.,

$$\mathbf{B}_s \sim \mathcal{N}(\mathbf{0}, \sigma_s^2\mathbf{I}), \tag{9}$$

ensuring multi-scale feature extraction across different frequency bands.

Using this transformation, the **temporal encoding matrix** $\mathbf{T} \in \mathbb{R}^{T\times Sd}$ is constructed as follows:

$$\mathbf{T} = \begin{bmatrix} \gamma(\tau_1) \\ \gamma(\tau_2) \\ \vdots \\ \gamma(\tau_T) \end{bmatrix} = \begin{bmatrix} \sin(2\pi B_1\tau_1) & \cos(2\pi B_1\tau_1) & \dots & \sin(2\pi B_S\tau_1) & \cos(2\pi B_S\tau_1) \\ \sin(2\pi B_1\tau_2) & \cos(2\pi B_1\tau_2) & \dots & \sin(2\pi B_S\tau_2) & \cos(2\pi B_S\tau_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sin(2\pi B_1\tau_T) & \cos(2\pi B_1\tau_T) & \dots & \sin(2\pi B_S\tau_T) & \cos(2\pi B_S\tau_T) \end{bmatrix}. \tag{10}$$

Using the temporal encoding matrix $\mathbf{T}$, the temporal dynamic matrix C_{temp} is shown as:

$$C_{temp} = \mathbf{T}\mathbf{T}^T \in \mathbb{R}^{T\times T}. \tag{11}$$

This formulation encodes dependencies across all time steps, naturally capturing periodicity and multi-scale temporal variations.

Regularization Constraint of Spatiotemporal Feature After obtaining the graph Laplacian matrix L_{rw} and the temporal dynamic matrixC_{temp}, they are incorporated into the Huber norm as the regularization constraints,

$$\begin{aligned} \varepsilon(P,E) = & \underbrace{\beta\|J\circ(W-PE)\|_{\text{Huber}}}_{\text{Data fitting term}} + \underbrace{\lambda_{reg}\left(\|P\|_F^2 + \|E\|_F^2\right)}_{\text{Tikhonov regularization}} \\ & + \underbrace{\lambda_1 \operatorname{trace}\left(P^T L_{rw} P\right)}_{\text{Spatial proximity constraint}} + \underbrace{\lambda_2 \operatorname{trace}\left(E^T C_{temp} E\right)}_{\text{Temporal dynamic constraint}} \end{aligned} \tag{12}$$

where λ_1 and λ_2 are two hyperparameters controlling the effects of spatial and temporal features, respectively. β is hyperparameter controlling the effects of the Huber norm.

3.2 Model Optimization

It is noteworthy that since the Huber-norm-based loss function in (12) exhibits gradient discontinuity at specific points, (12) can be reformulated into an element-wise representation, as shown in [11], [27], and [41].

$$
\begin{aligned}
\varepsilon(P,E) =& \beta \sum_{i=1}^{N}\sum_{j=1}^{T}\left(\begin{cases}\frac{1}{2}(w_{i,j}-p_{i,\cdot}e_{\cdot,j})^2, & \text{if } |w_{i,j}-p_{i,\cdot}e_{\cdot,j}|\le\delta\\ \delta(|w_{i,j}-p_{i,\cdot}e_{\cdot,j}|-\frac{1}{2}\delta), & \text{otherwise}\end{cases}\right)\\
&+\lambda_{\text{reg}}\sum_{i=1}^{N}\sum_{j=1}^{T}\left(\|p_{i,\cdot}\|_F^2+\|e_{\cdot,j}\|_F^2\right)+\lambda_1\sum_{i=1}^{N}\sum_{j=1}^{T}\left((P^TL_{rw}P)_{i,i}\right)\\
&+\lambda_2\sum_{i=1}^{N}\sum_{j=1}^{T}\left((E^TC_{temp}E)_{j,j}\right)
\end{aligned}
\tag{13}
$$

where $p_{i,\cdot}$ and $e_{\cdot,j}$ denote the ith row vector of P and the jth column vector of E, respectively. Then, considering the instant loss $\varepsilon_{i,j}$ of $\varepsilon(P,E)$ on a single entry $w_{i,j}$, we define

$$
\begin{aligned}
\varepsilon(i,j) =& \beta \left(\begin{cases}\frac{1}{2}(w_{i,j}-p_{i,\cdot}e_{\cdot,j})^2, & \text{if } |w_{i,j}-p_{i,\cdot}e_{\cdot,j}|\le\delta\\ \delta(|w_{i,j}-p_{i,\cdot}e_{\cdot,j}|-\frac{1}{2}\delta), & \text{otherwise}\end{cases}\right)\\
&+\lambda_{\text{reg}}\left((p_{i,\cdot})^2+(e_{\cdot,j})^2\right)+\lambda_1\sum_{m}(L_{rw})_{im}(p_{i,\cdot}-p_{m,\cdot})^2\\
&+\lambda_2\sum_{n}(C_{temp})_{jn}(e_{\cdot,j}-e_{\cdot,n})^2
\end{aligned}
\tag{14}
$$

The optimization with respect to top $p_{i,\cdot}$ and $e_{\cdot,j}$ can achieve it by the stochastic gradient descent (SGD) algorithm. Then, at the tth iteration, employ SGD to minimize (14) as follows:

$$
\begin{cases}
p_{i,\cdot}^{t}=p_{i,\cdot}^{t-1}-\eta\frac{\partial\varepsilon_{i,j}^{t-1}}{\partial p_{i,\cdot}^{t-1}}\\
e_{\cdot,j}^{t}=e_{\cdot,j}^{t-1}-\eta\frac{\partial\varepsilon_{i,j}^{t-1}}{\partial e_{\cdot,j}^{t-1}}
\end{cases}
\tag{15}
$$

where $p_{i,\cdot}^{t-1}$, $e_{\cdot,j}^{t-1}$, and $\varepsilon_{i,j}^{t-1}$ denote the states of $p_{i,\cdot}$, $e_{\cdot,j}$, and $\varepsilon_{i,j}$ at the $(t-1)$th iteration, and η denotes the learning rate of SGD. Let $\Delta_{i,j}^{t-1}=w_{i,j}-p_{i,\cdot}^{t-1}e_{\cdot,j}^{t-1}$ be the estimation error on a single entry $w_{i,j}$ at the $(t-1)$th iteration. By combining (14) into (15), the updating rules of $p_{i,\cdot}^{t}$ and $e_{\cdot,j}^{t}$ on a single entry $w_{i,j}$ at the tth iteration are obtained as follows:

$$
\Delta_{i,j}^{t-1}|\le\delta\begin{cases}
p_{ik}'=p_{ik}'^{-1}+\eta\beta\Delta_{i,j}^{t-1}e_{jk}^{t-1}-2\eta\lambda_{\text{reg}}p_{ik}^{t-1}\\
-2\eta\lambda_1\sum_m(L_{\text{rw}})_{im}(p_{ik}^{t-1}-p_{mk}^{t-1})\\
e_{jk}'=e_{jk}^{t-1}+\eta\beta\Delta_{i,j}^{t-1}p_{ik}^{t-1}-2\lambda_{\text{reg}}e_{jk}^{t-1}\\
-2\lambda_2\sum_n(C_{\text{temp}})_{jn}(e_{jk}^{t-1}-e_{nk}^{t-1})
\end{cases}
\tag{16}
$$

$$\Delta_{i,j}^{t-1}| > \delta \begin{cases} p_{ik}' = p_{ik}'^{-1} + \eta\beta\delta \cdot \mathrm{sgn}(\Delta_{i,j}^{t-1}) e_{jk}^{t-1} \\ -2\lambda_{\mathrm{reg}} p_{ik}^{t-1} - 2\lambda_1 \sum_m (L_{\mathrm{rw}})_{im} (p_{ik}^{t-1} - p_{mk}^{t-1}) \\ e_{jk}' = e_{jk}^{t-1} + \eta\beta\delta \cdot \mathrm{sgn}(\Delta_{i,j}^{t-1}) p_{ik}^{t-1} \\ -2\lambda_{\mathrm{reg}} e_{jk}^{t-1} - 2\lambda_2 \sum_n (C_{\mathrm{temp}})_{jn} (e_{jk}^{t-1} - e_{nk}^{t-1}) \end{cases} \tag{17}$$

where $\mathrm{sgn}(\cdot)$ is the sign function, defined as:

$$\mathrm{sgn}(x) = \begin{cases} 1, & x > 0 \\ -1, & x < 0 \end{cases} \tag{18}$$

3.3 Theoretical Analysis

Error Bound Analysis with Huber-Spatiotemporal Regularization

Proposition 1. *The composite loss at iteration t combines Huber loss and spatiotemporal regularization:*

$$O_t^{H,ST} = \beta_t^H \mathcal{H}(J \circ (W - P_t E_t)) + \beta_t^{ST} \left(\lambda_1 tr(P^T L_{rw} P) + \lambda_2 tr(E^T C_{temp} E) \right) \tag{19}$$

where $\mathcal{H}(\cdot)$ denotes Huber norm. The cumulative error $B_t^{H,ST} := \sum_{\omega=1}^t O_\omega^{H,ST}$ satisfies

$$B_T^{H,ST} \le \min B_T^H, B_T^{ST} + \ln 2\sqrt{\ln T} + \frac{T}{8\sqrt{\ln T}} \tag{20}$$

Proof. Define potential function $F_t = e^{-\sigma B_t^H} + e^{-\sigma B_t^{ST}}$ with $\sigma = (1/\ln T)^{1/2}$. Through Hoeffding's inequality:

$$\ln \frac{F_T}{F_0} \le -\sigma B_T^{H,ST} + \frac{\sigma^2 T}{8} \qquad \le -\sigma \min B_T^H, B_T^{ST} - \ln 2 \tag{21}$$

Rearranging terms yields the bound with $\sigma = \sqrt{1/\ln T}$.

Recovery Error Guarantee

Proposition 2. *For any corrupted matrix $H \in \mathbb{R}^{M\times N}$ satisfying $|H - \mathcal{P}(H)|_F \le \alpha |H|F$ $\alpha \in [0, 1)$, the reconstruction error satisfies:*

$$|X - W|F \le \frac{\lambda\sqrt{2rank(W)} + z_1|LrwW|F + z_2|WCtemp|F}{1 - \alpha} \tag{22}$$

where $\mathcal{P}(H) := \beta_1 J \circ sign(H) + \beta_2 J \circ H + z_1 LrwH + z_2 HCtemp$.

Proof. Apply KKT conditions to derive the reconstruction formula:

$$X = \mathrm{SVT}_\lambda(W - \mathcal{P}(W)) \tag{23}$$

Use triangle inequality with SVT properties:

$$|X - W|_F \le \lambda\sqrt{2\mathrm{rank}(W)} + |\mathcal{P}(W)|_F \tag{24}$$

Combine with spectral bounds for $\mathcal{P}(W)$.

Summary: The integration of Huber loss and spatiotemporal regularization ensures controlled cumulative errors ($\leq \min B_T^H, B_T^{ST} + \mathcal{O}(\sqrt{\ln T})$) and bounded reconstruction error proportional to noise level and graph/temporal smoothness.

Table 1: Properties of Experimental Datasets.

No.	Name	$\lvert\mathcal{M}\rvert$	$\lvert\mathcal{N}\rvert$	Time	Minimum/Maximum
D1	Beijing PM2.5 concentration [1]	35	8647	2014-05 – 2015-04	3.0 / 773.7
D2	Beijing CO Concentration[42]	35	8647	2014-05–2015-04	0.1 / 20.0
D3	Sea surface Temperature[24]	70	1733	1976-01–2014-12	0.01 / 30.31
D4	Daily mean CO concentration[35]	74	365	2010-01–2010-12	0.1 / 2.9

4 Experiments

The experiments are designed to address the following research questions (RQs):

RQ.1. Does the proposed ALFA-FE model outperform state-of-the-art models in recovering missing data in WSNs?

RQ.2. How do noise data affect the performance of ALFA-FE and other recovery models?

RQ.3. How do different hyperparameter settings influence the performance of the proposed ALFA-FE model?

4.1 Experimental Setup

Datasets. To evaluate the performance of the proposed model, we select four benchmark datasets collected from real-world WSNs, including PM2.5, CO, and sea surface temperature. The key properties of these datasets are summarized in Table 1.

Evaluation Metrics. Mean absolute error (MAE) and root mean square error (RMSE) are widely used to assess predictive accuracy [17, 23, 31]. RMSE penalizes larger errors due to the squared term, making it suitable for scenarios where extreme deviations matter, while MAE treats all errors uniformly, providing a balanced evaluation [25, 27, 20, 41, 22]. In WSN missing data recovery, where outliers are common, the metric play a crucial role. The mathematical formulations

Table 2: The comparison results of recovery accuracy on D1-D4.

Data	**Metric**	**ST-LRMA**	**BR-TVGS**	**LRDS**	**RRImpu**	**L3F**	**TRSS**	**LFA-STSR**	**ALFA-FE**
D1	MAE	12.3630±0.94	9.82±0.67	9.56±0.70	23.2691±1.09	13.8012±1.15	8.1215±1.50	8.2616±1.53	7.9241±1.54
	RMSE	22.7555±1.97	18.7477±1.01	18.0317±1.06	38.8918±1.50	25.35±1.04	16.2920±1.64	16.0696±1.64	15.4539±1.62
D2	MAE	0.1946±0.02	0.1658±0.02	0.1651±0.01	0.3027±0.02	0.2491±0.02	**0.1322**±0.02♦	0.1445±0.02	0.1393±0.01
	RMSE	0.4366±0.02	0.3777±0.01	0.3731±0.01	0.5507±0.02	0.5180±0.02	**0.3126**±0.02♦	0.3390±0.02	0.3220±0.02
D3	MAE	0.1268±0.02	0.1734±0.02	0.1260±0.01	0.4064±0.02	0.1255±0.02	0.1737±0.02	0.1207±0.0150	0.1162±0.02
	RMSE	0.1753±0.02	0.2438±0.01	0.1740±0.02	0.5449±0.01	0.1709±0.013	0.2444±0.03	0.1654±0.02	0.16±0.02
D4	MAE	0.1145±0.01	0.0981±0.01	0.0922±0.01	0.1064±0.01	0.0984±0.01	0.0919±0.01	0.0969±0.01	0.0780±0.01
	RMSE	0.1916±0.01	0.1585±0.01	0.1488±0.01	0.1667±0.01	0.1488±0.01	0.1478±0.01	0.1471±0.01	0.1124±0.01
Statistical	**win/loss**♦	8/0	8/0	8/0	8/0	8/0	6/2	8/0	—
	p-value⋆	0.002	0.002	0.002	0.002	0.002	0.002	0.002	—
Analysis	**F-rank**∗	5.78	5.11	4.11	8.00	7.00	2.33	2.67	**1.00**

are as follows:

$$
\begin{aligned}
\text{MAE} &= \frac{\sum\limits_{w_{i,j} \in \Gamma} |w_{i,j} - \hat{w}_{i,j}|}{|\Gamma|} \\
\text{RMSE} &= \sqrt{\frac{\sum\limits_{w_{i,j} \in \Gamma} (w_{i,j} - \hat{w}_{i,j})^2}{|\Gamma|}}
\end{aligned}
\tag{25}
$$

where $\hat{w}_{i,j}$ denotes the estimation of $w_{i,j}$ and Γ denotes the testing set, a lower value of RMSE/MAE indicates better accuracy.

Baselines. The proposed ALFA-FE model is evaluated against seven state-of-the-art models, including ST-LRMA[25], BR-TVGS[27], RRImpu[22], L^{3}F[41], TRSS[7], and LFA-STSR[40]. A brief description of the competing models is provided in supplement document.

Experimental Details. To simulate missing data scenarios in WSNs, partial observations are randomly selected from the complete dataset to construct the training set, while the remaining observations are designated as the testing set. Within the training set, half of the data is utilized for model training, and the other half is reserved for performance validation and hyperparameter tuning. After determining the optimal hyperparameters, model retraining is performed using the entire training set. The maximum number of iterations is fixed at 1500, and early termination is triggered if the error difference between two consecutive iterations falls below 10^{-6}. Each experiment is conducted five times, and the average results are reported. All experiments are executed on a computing system equipped with a 3.4-GHz Intel i7 processor and 64 GB of RAM.

4.2 Performance Comparison (RQ.1)

The recovery accuracy of ALFA-FE. Comprehensive evaluations at a 0.8 sampling rate demonstrate superior reconstruction capability of ALFA-FE across four datasets, as shown in Table 2. Statistically significant improvements in both accuracy and stability are achieved, with MAE=7.92±1.54 and RMSE=15.45±1.62

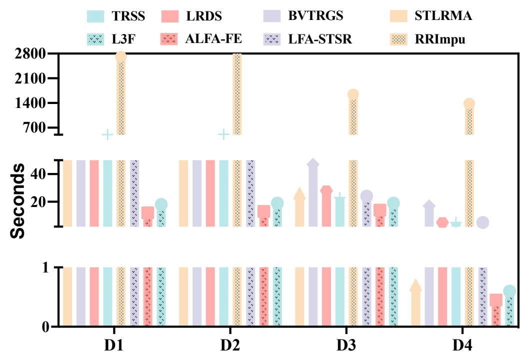


Fig. 2: The comparison CPU running time of involved models on D1–D4.

on D1, surpassing LFA-STSR by 4.3% and 3.9%, respectively. On D4, MAE is reduced to 0.078±0.006, marking a 15.1% improvement over TRSS (0.0919±0.007). In D2, variance remains 50% lower than TRSS (0.139±0.01 vs. 0.132±0.02) while maintaining comparable accuracy. For temporal reconstruction in D3, RMSE reaches 0.1569±0.0154, with 11.2% error reduction compared to LRDS (0.174±0.0149). Dominance in pairwise comparisons is evident with 8/0 win-loss ratio, significant differentiation confirmed by Wilcoxon tests ($p<0.002$), and optimal Friedman ranking (F-rank=1.00 vs. TRSS=2.33). Despite slightly better D2 metrics (MAE=0.132 vs. 0.139), TRSS exhibits compromised robustness, reflected in 100% higher variance and inferior ranking. This systematic validation highlights the effectiveness of ALFA-FE in balancing precision-stability trade-offs through adaptive spatiotemporal regularization. Compared to TRSS, ALFA-FE yields the largest relative gain on D1 and D4, where sharp pollutant spikes are prevalent, indicating the benefit of Huber-based anomaly preservation. On smoother datasets like D3, the improvements are moderate, demonstrating Fourier embeddings' capacity to align with continuous temporal structures.

Comparison of Computational Efficiency. To evaluate the computational efficiency of all test models, we measured their CPU runtime across all datasets, as illustrated in Fig. 2. Besides, Analysis of computational complexity can be found in the supplementary document. In Our model generally achieves a lower CPU runtime compared to the baseline models, except in certain cases where L3F performs better. This is primarily due to the fact that our approach incorporates spatio-temporal smoothness characteristics, which are not considered in the L3F model.

4.3 Robustness Analysis under Noisy Conditions (RQ.2)

To rigorously evaluate the capability of ALFA-FE in handling heterogeneous anomalies and coupled noise-outlier interactions inherent in real-world WSN deployments, we design a structured noise injection protocol that emulates the

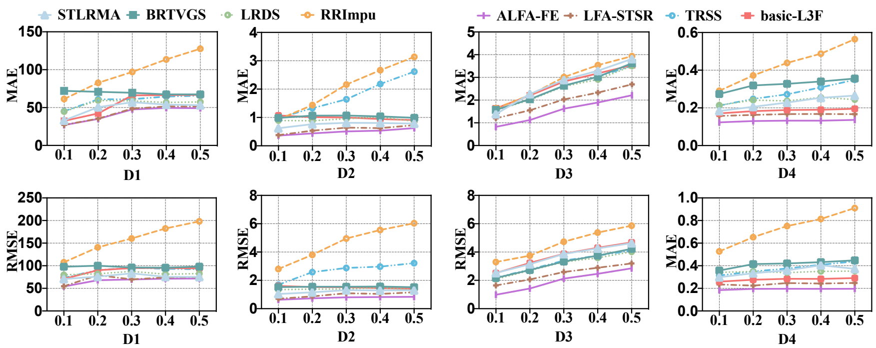


Fig. 3: Outlier data sensitivity tests on all the datasets.

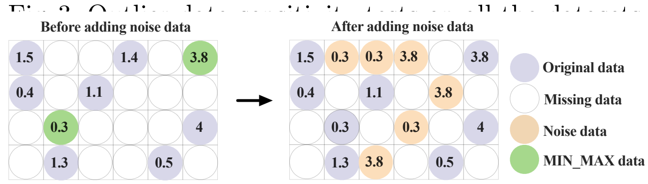


Fig. 4: An example of adding noise data.

spatiotemporal characteristics of sensor data corruption. The contamination process adheres to four principles: (a) Spatially randomized selection of observed entries to simulate uneven anomaly distribution across different district sensor arrays; (b) Value corruption via extremal substitution based on localized training subsets, replicating multimodal WSNs data distributions observed in Fig.1b; (c) Controlled escalation of noise intensity from 10% to 50% contamination ratios to mimic progressive sensor degradation scenarios; (d) Strict isolation of noise injection to training data, preserving test set integrity for unbiased generalization assessment. As visualized in Fig.4, our simulation framework preserves the baseline distribution (pink) while introducing context-aware outliers (orange) bounded by localized extremal thresholds (green). Quantitative analysis in Fig. 3 reveals that conventional low-rank models exhibit accelerated performance decay under rising contamination levels due to undifferentiated noise-anomaly processing. In contrast, our ALFA-FE framework demonstrates moderated accuracy degradation, maintaining 22% lower MAE than the best baseline at 50% noise ratio. The results confirm the dual capacity of ALFA-FE that is simultaneously recovering stable low-rank structures through Fourier-embedded manifolds and protecting sensor-specific anomaly signatures via noise-outlier discriminative learning.

4.4 Effects by the Hyperparameters (RQ.3)

This section investigates the influence of *spatio-temporal regularization parameters* (λ_1 and λ_2) and the *latent factor dimension* (k) on model effectiveness. We

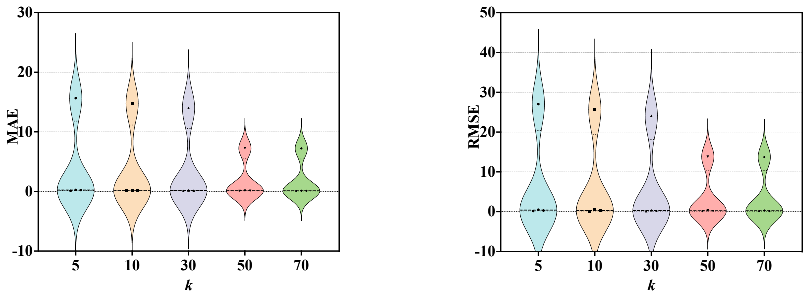


Fig. 5: The MAE/RMSE of ALFA-FE as k increases from 5 to 70 on all the datasets.

systematically analyze their effects on reconstruction accuracy, stability, and adaptability to varying data patterns.

Trade-off Analysis of Feature Dimensions. Increasing the feature dimension generally improves model recovery accuracy at the expense of increased computational time. At sampling rate of 0.9, we experimented with feature dimensions ranging from 5 to 80, and the results are presented in Fig. 5. The data indicate that both MAE and RMSE decrease as the feature dimension increases, until reaching a point beyond which further improvements become marginal. Furthermore, as analysised in Section 3.3 and corroborated by the violin plots in the accompanying figures, increasing the parameter k leads to a notable rise in time complexity. Therefore, a value of $k = 50$ is recommended to strike a balance between recovery accuracy and computational efficiency.

Impact of Spatiotemporal Regularization on Reconstruction Performance. Fig. 6 illustrates the impact of spatiotemporal regularization on reconstruction performance under sampling rate of 0.1. A systematic parameter analysis reveals non-linear error responses to parameter variations: Incremental increases in λ_1 and λ_2 (ranging from 0.2 to 1.0) initially lead to a 14.5% reduction in MAE on dataset D3. However, beyond a critical threshold, further parameter augmentation results in performance deterioration. Dataset D4 achieves optimal performance at $\lambda_1 = 0.02$ and $\lambda_2 = 0.005$, yielding a minimum MAE of 0.1173. This result underscores the importance of balancing spatial and temporal regularization constraints to achieve optimal reconstruction accuracy. The interplay between spatial λ_1 and temporal λ_2 regularization parameters exhibits a complementary effect, where appropriately tuned values mitigate reconstruction errors while preserving structural integrity. Overall, harmonized spatiotemporal regularization plays a pivotal role under conditions of limited sampling.

5 Conclusion

This article proposes ALFA-FE for robust spatiotemporal signal recovery in wireless sensor networks. Its main idea is twofold: 1) embedding Concatenated Fourier

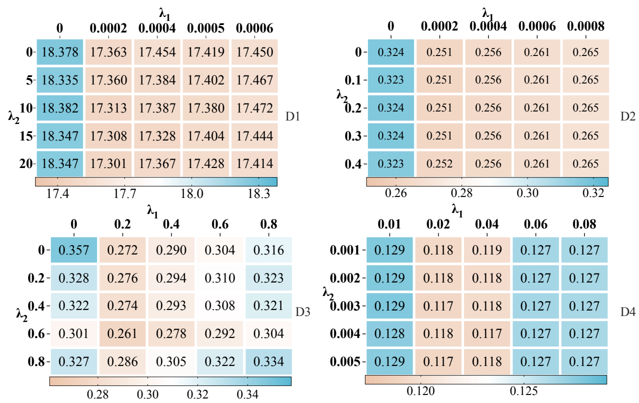


Fig. 6: The MAE of ALFA-FE with different λ_1 and λ_2 on all the datasets.

Features into temporal dynamics to improve the adaptability to non-stationary data patterns; 2) integrating the Huber norm with gradient descent optimization to enhance robustness against coupled noise and anomalies. By harmonizing frequency-domain regularization and adaptive robustness constraints, ALFA-FE possesses the merits of both dynamic representation learning and effective anomaly handling. In the experiments, ALFA-FE is evaluated on four benchmark datasets of varying scales from real-world WSNs. The results demonstrate that ALFA-FE significantly outperforms state-of-the-art methods in reconstruction accuracy and robustness under noisy conditions. Although ALFA-FE performs well, its deployment can be streamlined. Replacing fixed-rate SGD and grid search with adaptive optimizers (e.g., Adam) and automated tuning (e.g., Bayesian optimization) will cut training cost, while extending the framework to edge-friendly, decentralized updates and dynamic graphs will let it handle mobile or wireless WSNs and meet real-time, resource-limited constraints in large-scale IoT deployments.

References

1. Beijing Municipal Ecological and Environmental Monitoring Center: Beijing $\mathrm{pm}_{2.5}$ data (2016), https://quotsoft.net/air/, [Online]
2. Che, H., Pu, X., Ouyang, D., Li, B.: Enhanced tensorial self-representation subspace learning for incomplete multi-view clustering. In: Proceedings of the 32nd ACM International Conference on Multimedia. pp. 719–728 (2024)
3. Chen, Y., Chi, Y., Goldsmith, A.J.: Exact and stable covariance estimation from quadratic sampling via convex programming. IEEE Transactions on Information Theory **61**(7), 4034–4059 (2015)
4. Dowlatshahi, M.B., Rafsanjani, M.K., Gupta, B.B.: An energy aware grouping memetic algorithm to schedule the sensing activity in wsns-based iot for smart cities. Applied Soft Computing **108**, 107473 (2021)
5. Fadhel, M.A., Duhaim, A.M., Saihood, A., Sewify, A., Al-Hamadani, M.N., Albahri, A., Alzubaidi, L., Gupta, A., Mirjalili, S., Gu, Y.: Comprehensive systematic review

of information fusion methods in smart cities and urban environments. Information Fusion p. 102317 (2024)
6. Ghazal, T.M., Hasan, M.K., Alzoubi, H.M., Alshurideh, M., Ahmad, M., Akbar, S.S.: Internet of things connected wireless sensor networks for smart cities. In: The effect of information technology on business and marketing intelligence systems, pp. 1953–1968. Springer (2023)
7. Giraldo, J.H., Mahmood, A., Garcia-Garcia, B., Thanou, D., Bouwmans, T.: Reconstruction of time-varying graph signalsvia sobolev smoothness. IEEE Transactions on Signal and Information Processing over Networks **8**, 201–214 (2022)
8. Gong, X., Yu, L., Wang, J., Zhang, K., Bai, X., Pal, N.R.: Unsupervised feature selection via adaptive autoencoder with redundancy control. Neural Networks **150**, 87–101 (2022)
9. Goswami, P., Mukherjee, A., Sarkar, B., Yang, L.: Multi-agent-based smart power management for remote health monitoring. Neural Computing and Applications **35**(31), 22771–22780 (2023)
10. Grover, A., Lall, B.: A recursive method for estimating missing data in spatio-temporal applications. IEEE Transactions on Industrial Informatics **18**(4), 2714–2723 (2021)
11. He, X., Tang, J., Du, X., Hong, R., Ren, T., Chua, T.S.: Fast matrix factorization with nonuniform weights on missing data. IEEE transactions on neural networks and learning systems **31**(8), 2791–2804 (2019)
12. Kong, X., Wang, J., Hu, Z., He, Y., Zhao, X., Shen, G.: Mobile trajectory anomaly detection: Taxonomy, methodology, challenges, and directions. IEEE Internet of Things Journal (2024)
13. Li, R., Zhou, G., Li, X., Jia, L., Shang, Z.: Semi-supervised multi-label dimensionality reduction learning based on minimizing redundant correlation of specific and common features. Knowledge-Based Systems **294**, 111789 (2024)
14. Liu, F., Zhang, A., Du, R., Sun, G.: Structural sparse tensor robust pca method based on cross-entropy for vehicular network data recovery. IEEE Transactions on Vehicular Technology (2024)
15. Liu, H., Wang, P., Huang, L., Qu, Q., Balzano, L.: Symmetric matrix completion with relu sampling. In: Proceedings of the 41st International Conference on Machine Learning. ICML'24, JMLR.org (2024)
16. Long, J., Wu, H.: Latent factor analysis enhanced graph contrastive learning for recommendation. In: 2024 IEEE International Conference on Systems, Man, and Cybernetics (SMC). pp. 547–552. IEEE (2024)
17. Luo, X., Wu, H., Wang, Z., Wang, J., Meng, D.: A novel approach to large-scale dynamically weighted directed network representation. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(12), 9756–9773 (2021)
18. Luo, X., Yuan, Y., Chen, S., Zeng, N., Wang, Z.: Position-transitional particle swarm optimization-incorporated latent factor analysis. IEEE Transactions on Knowledge and Data Engineering **34**(8), 3958–3970 (2020)
19. Majid, M., Habib, S., Javed, A.R., Rizwan, M., Srivastava, G., Gadekallu, T.R., Lin, J.C.W.: Applications of wireless sensor networks and internet of things frameworks in the industry revolution 4.0: A systematic literature review. Sensors **22**(6), 2087 (2022)
20. Mao, X., Qiu, K., Li, T., Gu, Y.: Spatio-temporal signal recovery based on low rank and differential smoothness. IEEE Transactions on Signal Processing **66**(23), 6281–6296 (2018)
21. Miao, J., Kou, K.I.: Color image recovery using low-rank quaternion matrix completion algorithm. IEEE Transactions on Image Processing **31**, 190–201 (2021)

22. Muzellec, B., Josse, J., Boyer, C., Cuturi, M.: Missing data imputation using optimal transport. In: International Conference on Machine Learning. pp. 7130–7140. PMLR (2020)
23. Pan, J., Gillis, N.: Generalized separable nonnegative matrix factorization. IEEE transactions on pattern analysis and machine intelligence **43**(5), 1546–1561 (2019)
24. Physical Sciences Division, Earth System Research Laboratory, National Oceanic and Atmospheric Administration: Sea surface temperature (sst) v2 (2015), http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html, accessed: Dec. 18, 2015. [Online]
25. Piao, X., Hu, Y., Sun, Y., Yin, B., Gao, J.: Correlated spatio-temporal data collection in wireless sensor networks based on low rank matrix approximation and optimized node sampling. Sensors **14**(12), 23137–23158 (2014)
26. Qin, W., Luo, X., Li, S., Zhou, M.: Parallel adaptive stochastic gradient descent algorithms for latent factor analysis of high-dimensional and incomplete industrial data. IEEE Transactions on Automation Science and Engineering (2023)
27. Qiu, K., Mao, X., Shen, X., Wang, X., Li, T., Gu, Y.: Time-varying graph signal reconstruction. IEEE Journal of Selected Topics in Signal Processing **11**(6), 870–883 (2017)
28. Rashid, B., Rehmani, M.H.: Applications of wireless sensor networks for urban areas: A survey. Journal of network and computer applications **60**, 192–219 (2016)
29. Sarkar, S., Sahay, R.R.: A non-local superpatch-based algorithm exploiting low rank prior for restoration of hyperspectral images. IEEE Transactions on Image Processing **30**, 6335–6348 (2021)
30. Sharma, H., Haque, A., Blaabjerg, F.: Machine learning in wireless sensor networks for smart cities: a survey. Electronics **10**(9), 1012 (2021)
31. Shi, W., Zhu, Y., Philip, S.Y., Zhang, J., Huang, T., Wang, C., Chen, Y.: Effective prediction of missing data on apache spark over multivariable time series. IEEE Transactions on Big Data **4**(4), 473–486 (2017)
32. Soares, C.A.R., de Souza Couto, R., Sztajnberg, A., do Amaral, J.L.M.: Posimnet-r: An immunologic resilient approach to position routers in industrial wireless sensor networks. Expert Systems with Applications **188**, 116045 (2022)
33. Song, J., Xia, S., Wang, J., Patel, M., Chen, D.: Uncertainty quantification of hyperspectral image denoising frameworks based on sliding-window low-rank matrix approximation. IEEE Transactions on Geoscience and Remote Sensing **60**, 1–12 (2021)
34. Swami Durai, S.K., Duraisamy, B., Thirukrishna, J.: Certain investigation on healthcare monitoring for enhancing data transmission in wsn. International journal of wireless information networks **30**(1), 103–110 (2023)
35. United States Environmental Protection Agency: Air quality data (May 2016), https://www.epa.gov/outdoor-air-quality-data, [Online]
36. Wang, F., Xu, Z., Chen, J., Hu, R.: Semisupervised action recognition with adaptive correlation learning. International Journal of Machine Learning and Cybernetics pp. 1–14 (2024)
37. Wang, J., Han, H., Li, H., He, S., Sharma, P.K., Chen, L.: Multiple strategies differential privacy on sparse tensor factorization for network traffic analysis in 5g. IEEE transactions on industrial informatics **18**(3), 1939–1948 (2021)
38. Wang, J., Ma, Z., Nie, F., Li, X.: Top-k discriminative feature selection with uncorrelated and $\ell 2,0$-norm equation constraints. Neurocomputing **598**, 128069 (2024)
39. Wang, Z.Y., So, H.C., Zoubir, A.M.: Robust low-rank matrix completion via sparsity-inducing regularizer. Signal Processing **226**, 109666 (2025)

40. Wu, D., Li, Z., Yu, Z., He, Y., Luo, X.: Robust low-rank latent feature analysis for spatiotemporal signal recovery. IEEE Transactions on Neural Networks and Learning Systems (2023)
41. Wu, D., Shang, M., Luo, X., Wang, Z.: An l 1-and-l 2-norm-oriented latent factor model for recommender systems. IEEE Transactions on Neural Networks and Learning Systems **33**(10), 5775–5788 (2021)
42. Xu, P., Ruan, W., Sheng, Q.Z., Gu, T., Yao, L.: Interpolating the missing values for multi-dimensional spatial-temporal sensor data: A tensor svd approach. In: Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services. p. 442–451. MobiQuitous 2017, Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3144457.3144474, https://doi.org/10.1145/3144457.3144474
43. Xu, Z., Zhang, Y., Ma, C., Yan, Y., Peng, Z., Xie, S., Wu, S., Yang, X.: Lere: learning-based low-rank matrix recovery with rank estimation. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 38, pp. 16228–16236 (2024)
44. Zhang, H., Huang, T.Z., Zhao, X.L., Zhang, S., Xie, J.Y., Jiang, T.X., Ng, M.K.: Learnable transform-assisted tensor decomposition for spatio-irregular multidimensional data recovery. ACM Transactions on Knowledge Discovery from Data **19**(1), 1–23 (2024)
45. Zhang, H., Wu, M., Feng, Q., Li, H.: Aerqp: adaptive embedding representation-based qos prediction for web service recommendation. The Journal of Supercomputing **80**(3), 3042–3065 (2024)
46. Zhang, R., Qi, L., Yan, C., Chen, Z., Gong, W., Xu, Y., Rafique, W.: Privacy-aware web apis recommendation for consumer mashup creation based on iterative quantification. IEEE Transactions on Consumer Electronics (2024)