

MapT-STC: A One-Shot Spatiotemporal Data Completion Framework for Cold-Start Tasks in Sparse Crowdsensing

Mijia Zhang[✉], Member, IEEE, Wei Li[✉], En Wang[✉], Member, IEEE, and Jie Wu[✉], Fellow, IEEE

Abstract—Sparse crowdsensing tackles data incompleteness through sparse sensing coupled with inference-based completion. Yet prevailing solutions face three inherent constraints: (i) Spatiotemporal completion methods demand extensive historical/cross-domain data for source-domain pretraining and target-domain fine-tuning; (ii) Such dependency causes catastrophic performance degradation in cold-start scenarios lacking target-domain samples, a setting that remains largely unaddressed by existing data-dependent frameworks; (iii) Resource-intensive pretraining-fine-tuning pipelines hinder lightweight deployment. To overcome these challenges, this paper proposes Mapping Transfer Learning-based Spatiotemporal Completion method (MapT-STC), a one-shot spatiotemporal completion framework that advances beyond few-shot approaches Correlated Data Fusion for Matrix Completion (CDFMC) by featuring a remodeled pretraining-finetuning paradigm. Departing from conventional approaches, MapT-STC implements a dual-path learning system: it simultaneously trains a target-specialized model for domain-specific features and a cross-domain generalizable model via Correlation Alignment (CORAL). The framework explicitly aligns feature distributions by decomposing spatiotemporal data into separable temporal patterns and spatial structures, thereby enhancing the quality of domain adaptation. Ultimately, Kalman-filter-driven fusion integrates predictions from both specialized and generalized models for optimal completion. Experiments demonstrate MapT-STC's superiority under cold-start conditions: reducing initial error by over 30% and accelerating convergence time by at least 7.4% compared to state-of-the-art baselines, using only one target-domain training sample.

Index Terms—Sparse crowdsensing, one-shot spatiotemporal completion, dual-path learning, Kalman-filter fusion.

Received 15 June 2025; revised 25 February 2026; accepted 24 March 2026. Date of publication 31 March 2026; date of current version 6 August 2026. This work was supported in part by the National Natural Science Foundation of China under Grant 92567204, in part by the Jilin Province Major Science and Technology Project under Grant 20240212003GX, and in part by the Postdoctoral Fellowship Program of CPSF under Grant GZC20251096. An earlier version of this paper was presented at the IEEE INFOCOM 2024-IEEE Conference on Computer Communications [DOI: 10.1109/INFOCOM52122.2024.10621156]. Recommended for acceptance by R. Zhang. (Corresponding author: Wei Li.)

Mijia Zhang is with the College of Computer Science and Technology, Jilin University, Changchun 130012, China (e-mail: zhangmijia@jlu.edu.cn).

Wei Li is with the School of Business and Management, Jilin University, Changchun 130012, China (e-mail: W_li@jlu.edu.cn).

En Wang is with the College of Computer Science and Technology, Jilin University, Changchun 130012, China (e-mail: wangen@jlu.edu.cn).

Jie Wu is with the China Telecom Cloud Computing Research Institute, Beijing 100088, China, and also with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA 19122 USA (e-mail: jiewu@temple.edu).

Digital Object Identifier 10.1109/TMC.2026.3678980

I. INTRODUCTION

THE recent convergence of 5G mobile networks [1] and pervasive computing infrastructures [2] has catalyzed transformative advances in distributed sensing paradigms, enabling the emergence of Mobile Crowdsensing (MCS) as a disruptive data acquisition framework [3], [4], [5], [6], [7]. This methodology leverages decentralized participant networks to collect heterogeneous, large-scale datasets from diverse sources—including mobile devices, IoT sensors, and human contributors—while employing advanced machine learning algorithms for data processing and knowledge extraction [8], [9], [10]. The deployment efficacy of MCS has been demonstrated across multifarious domains such as real-time traffic flow prediction and management [11], [12], high-resolution environmental monitoring [13], and personalized healthcare delivery [14]. Nevertheless, critical limitations persist in scenarios characterized by limited device density or low user participation, wherein traditional MCS techniques exhibit substantial performance degradation due to spatial-temporal data sparsity. To address this fundamental challenge, sparse crowdsensing [15] has emerged as a strategic solution, utilizing inference-driven data completion mechanisms to augment service quality under constrained sensing coverage [16]. Crucially, however, the one-shot cold-start data completion problem, wherein models must handle entirely novel sensing tasks with minimal to no target, domain training data—remains unresolved for existing Big Data-dependent frameworks. This creates a fundamental barrier for deploying sensing systems in new domains where historical data collection is infeasible. As empirically characterized in the upper part of Fig. 1.

Fig. 1 delineates an end-to-end workflow for spatiotemporal data completion. The artificial segmentation of urban terrain into 4×4 subareas establishes a data collection paradigm wherein stochastic participant mobility and resource constraints preclude complete coverage of all 16 grids during any observational interval. In other words, urban maps are spatially tokenized into grid representations (e.g., timestamps 7:30, 8:00, 8:30), where sparse inputs are collected under real-world constraints; the Data Completion Neural Network processes these fragmented vectors into dense outputs; error computing (exemplified by MSE) evaluates performance, with the ideal trajectory (orange “Ideal” curve) targeting near-one initial error and convex decay convergence (≤ 75 iterations to stabilization), sharply

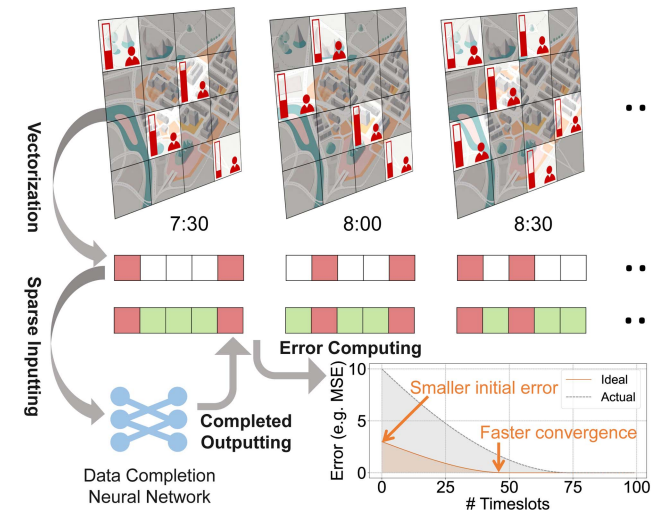


Fig. 1. Exemplary scenarios illustrating one-shot data completion challenges in sparse crowdsensing.

outperforming the sluggish actual baseline (gray “Actual” curve) as signified by the “Faster convergence” annotation. This problem can create severe consequences, such as hindering the observation of the environmental state of the forest. If we cannot obtain the complete temperature, humidity, and air quality data in time, it may lead to the fire not being detected in time and then cause the fire to spread in the forest. Despite the common occurrence of this problem in practical applications, inferring small amounts of initial data while facing the sparse crowdsensing has the following three distinct challenges: (i) *the ambiguity in spatiotemporal relationships*, (ii) *the reliance on data from a large number of historical timeslots*, and (iii) *the sparsity of the sensed data*.

To ground our discussion in a concrete application scenario, consider the challenge of monitoring air quality (e.g., $PM_{2.5}$) in a new city (the target domain) with very few sensors. A highly relevant but distinct dataset, such as humidity readings from a well-established sensor network in another city (the source domain), may exhibit strong correlations with the target $PM_{2.5}$ data, as conceptually illustrated in Fig. 2. The fundamental challenge we address is how to leverage these inter-data correlations, which can be positive or negative, to accurately infer missing data in the target domain from the very beginning, without requiring extensive target-specific data for training.

In practical deployments, sparse crowdsensing for novel tasks often suffers from contributor scarcity. This severely constrains usable information for inferring unsensed data, primarily due to inherent data sparsity and one-shot conditions. Taking the environmental monitoring systems as an example: in the initial stages of training the data sensing model, there may be limited data available due to the lack of historical data. Such constraints impede environmental adaptation, workflow implementation, and latent threat identification. Thus, knowledge transfer leveraging relevant prior crowdsensing tasks is essential. (e.g. the historical data of a similar environment in another city).

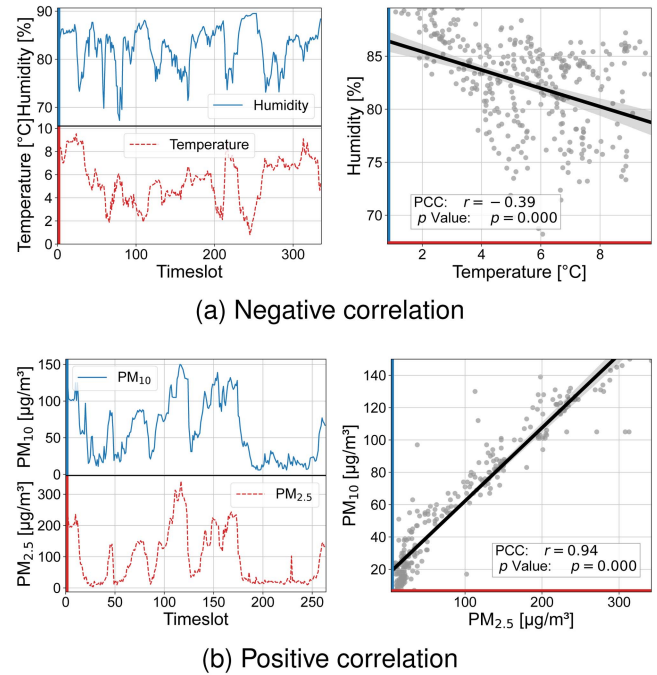


Fig. 2. Two exemplary manifestations of cross-dataset correlations between source and target domains.

However, identifying valuable tasks for knowledge acquisition and devising effective mechanisms for their utilization remain significant challenges. It shows the difficulty of obtaining additional information. Therefore, how to *efficiently obtain and exploit cross-task knowledge under one-shot conditions* is the first challenge.

As a matter of fact, one-shot data completion for novel sparse crowdsensing tasks remains imperfectly resolved. While existing inference methods—including VAE [17], DMF [18], [19], and IGMCM [20]—exhibit heavy reliance on massive sensed data, they typically suffer from suboptimal accuracy and elevated error rates in one-shot scenarios. It is imperative to harness both intra-data structures and inter-data correlations, integrating them with ML-driven computing paradigms such as Transfer Learning [21] and Domain Generalization [22], [23] to enable robust unsensed data inference. Whereas transfer learning presumes domain invariance, domain generalization operates under the premise of cross-domain diversity. To eliminate temporal dependencies, it is essential to holistically reconcile these opposing characteristics and dynamically establish equilibrium between similarity preservation and diversity adaptation. This complexity escalates significantly. Accordingly and consequently, how to *dynamically optimize the domain transfer to eliminate the number of target temporal data dependencies* is the second challenge.

Effectively leveraging intra- and inter-data correlations is of critical importance to seek solutions for addressing challenges in sparse crowdsensing. Evidence from Fig. 2 illustrates two pairs of datasets that demonstrate positive and negative correlations, respectively. Such robust negative or positive correlations are typically present in real-world datasets, making data from other

relevant datasets a valuable resource to tackle sparse crowdsensing tasks.

Prior research reveals a significant limitation: prevailing methods [19], [20], [24] predominantly address intra-data features while neglecting inter-data correlations. Although pioneering work like [16] incorporates both dimensions for sparse crowdsensing, it confines cross-dataset relationship measurement solely to the Pearson Correlation Coefficient (PCC). In most cases, it is reasonable to use the value of the PCC, but it also makes the algorithm unable to adjust dynamically. For sparse sensed data, especially in a cold-start task, it is difficult to correlate multiple datasets by calculating the PCC. We must find a new way to construct a spatiotemporal feature association between the target dataset and the source dataset, so as to reduce the model fine-tuning cost. Therefore, how to **construct adaptive cross-dataset spatiotemporal couplings to replace static PCC metrics, thereby enabling effective knowledge transfer for the completion task under one-shot conditions**, is the third challenge.

To address the challenges above, while the previous work CDFMC [25] tackled few-shot data completion by leveraging auxiliary datasets, this journal version advances the field by introducing MapT-STC, a one-shot framework that operates under cold-start conditions. Key innovations include: (i) a mapping transfer learning approach via CORAL to generate pseudo-historical data without target-domain samples, (ii) a remodeled dual-path learning system that dynamically balances domain-specific and generalizable models, and (iii) formal guarantees for cross-domain feature transferability. These contributions fundamentally reduce reliance on historical timeslots and enable lightweight deployment. The proposed MapT-STC model thus establishes a novel paradigm for spatiotemporal matrix completion, effectively harnessing cross-source correlation synergies to elevate reconstruction precision and operational robustness. Unlike zero-shot approaches that require no target data, we focus on one-shot cold-start tasks, where models must handle novel sensing scenarios with only a minimal number of initial samples (e.g., a single timeslot), extending beyond few-shot settings and leveraging source-domain pretraining to enable chain-like propagation of inferences. Its extensible architecture demonstrates significant potential for applications transcending sparse crowdsensing, including fundamental data science and trustworthy AI systems.

Core innovations and technical contributions of this work are summarized as follows:

- We address the one-shot spatiotemporal data completion problem in cold-start sparse crowdsensing with MapT-STC, a robust framework that models uncertainty propagation and handles heterogeneous noise.
- By integrating crowdsensing with transfer learning and domain generalization, MapT-STC uses tensor decomposition, CORAL-based pseudo-history, and adaptive fusion to extract latent dependencies from sparse inputs.
- Unlike correlation-based methods, our framework offers theoretical guarantees for cross-source feature transfer under sparsity and manages bias-variance tradeoffs in noisy

settings, achieving SOTA accuracy with lower fine-tuning costs.

- MapT-STC establishes a data-efficient learning paradigm beyond crowdsensing, proving effective in high-stakes, data-scarce environments and laying groundwork for trustworthy AI applications.

To establish a logically coherent framework for this manuscript, we have systematically organized the subsequent content into the following structurally defined sections. In Section II, we will explore related research in the field. Next, in Section III, we will present the system model and the specific problem that we aim to solve. Sections IV and V will provide a comprehensive overview of our MapT-STC, including its intricate details. In Section VI, we will conduct thorough simulations to evaluate the performance of our approach. Finally, we will summarize our findings and conclusions in Section VII.

II. RELATED WORK

A. Sparse Crowdsensing

Crowdsensing technology emerges as an innovative paradigm, leveraging ubiquitous smart devices carried by mobile users to execute diverse sensing tasks, as substantiated by foundational studies [6] and [26]. It mainly studies core issues such as task allocation [27] and path planning [28]. Lin et al. [27] proposes an aggregation-based dual-heterogeneous allocation algorithm to address the NP-hard challenge of task allocation in spatial crowdsensing with coexisting heterogeneous tasks and workers. Zhu et al. [28] proposes DLMV, a scheme that predicts vehicle trajectories to optimize the selection of vehicles within a limited budget for mobile crowdsensing, maximizing sensing data collection (e.g., environmental monitoring, traffic noise monitoring, etc.) in urban application areas. However, unrestricted participant recruitment is often constrained by limited sensing budgets [29]. Crowdsensing systems face fundamental cost constraints in balancing participant incentives, resource consumption, and desired spatial-temporal coverage while maintaining data quality requirements. To address these problems, strategically selected sparse sampling can achieve comparable sensing fidelity by exploiting spatiotemporal correlations across subareas, as extensively validated in [30]. By leveraging spatiotemporal correlations extracted from historical datasets, unsensed areas can be probabilistically inferred through statistical inference frameworks, as demonstrated in the seminal work of Xie et al. [31]. Recent years have witnessed a significant paradigm shift, with researchers increasingly dedicated to advancing sparse crowdsensing frameworks that strategically sample limited subareas while reconstructing unsensed regions through computational inference techniques. For instance, Chen et al. [32] introduces HMSS to accurately estimate fine-grained air pollution maps from sparse, dynamically changing data collected by non-dedicated vehicles. It dynamically weights the physics model (effective with sparse data) and data-driven model (effective with dense data) based on data coverage and uncertainty. Wang et al. [33] develops a privacy-preserving framework for sparse crowdsensing. It generates an optimal

obfuscation matrix and adjusts the sensing data to fit the obfuscated location. Consequently, these established implementations of sparse crowdsensing provide a foundation for future methodological innovation, demonstrating significant potential to drive advancements in urban sensing paradigms.

B. Spatiotemporal Data Completion

From a mathematical modeling perspective, sparse crowdsensing can be conceptualized as completing a sparse spatiotemporal matrix. In practical applications, such matrices exhibit an inherent low-rank structure due to spatiotemporal correlations between adjacent timeslots and subareas. While Matrix Factorization (MF) has proven effective for inferring linearly correlated data, its efficacy is constrained when handling nonlinear spatiotemporal patterns. To address this, Fan and Cheng [18] introduced Deep Matrix Factorization (DMF) incorporating nonlinear neural networks. Extending DMF, Wang et al. [19] developed an end-to-end (E2E) model enabling concurrent sensing and high-precision future prediction for industrial scenarios. Furthermore, Zhang et al. [20] proposed Inductive Graph-based Matrix Completion (IGMC) leveraging Graph Neural Networks (GNN), outperforming DMF on specific datasets. Yoon et al. [34] proposes the Generative Adversarial Imputation Nets (GAIN), which is a novel method that adapts the Generative Adversarial Networks (GAN) framework for missing data imputation. It uses a generator to impute missing values conditioned on observed data, a discriminator to distinguish observed/imputed components, and introduces a hint mechanism to ensure the generator learns the true data distribution. Notably, Wang et al. [24] enhanced robustness by integrating outlier resilience mechanisms, while Xie et al. [35] designed a two-phase matrix completion framework exploiting environmental data features for fractured spatiotemporal data recovery.

The pervasive deployment of Big Data analytics has catalyzed extensive research into intra- and inter-data correlations across diverse domains. Wang et al. [16] pioneered multi-task correlation analysis through rigorous statistical verification of dominant correlation coefficients. Complementarily, Liu et al. [36] designed a multi-branch network architecture extracting intra-task features independently before concatenating outputs to model inter-task dependencies. Practical implementations includemobile crowdsensing with differential and distortion location Han et al.'s [37] multi-adversarial spatiotemporal GNN for meteorological forecasting, and critically in sparse crowdsensing, Wang et al. [38] leveraged cross-modal data interactions to enhance outlier detection robustness. While Pearson Correlation Coefficient (PCC) remains prevalent for cross-dataset analysis, conventional machine learning models exhibit catastrophic failure under out-of-distribution (OOD) scenarios. This limitation motivated domain generalization research [39], [40], [41], where predictors trained on related domains generalize to unseen environments. Robey et al. [23] developed non-convex duality-based algorithms with bounded duality gaps and convergence guarantees. Wang et al. [22] systematically categorized methodologies into data manipulation, representation learning, and strategic optimization. Contrastingly, transfer learning

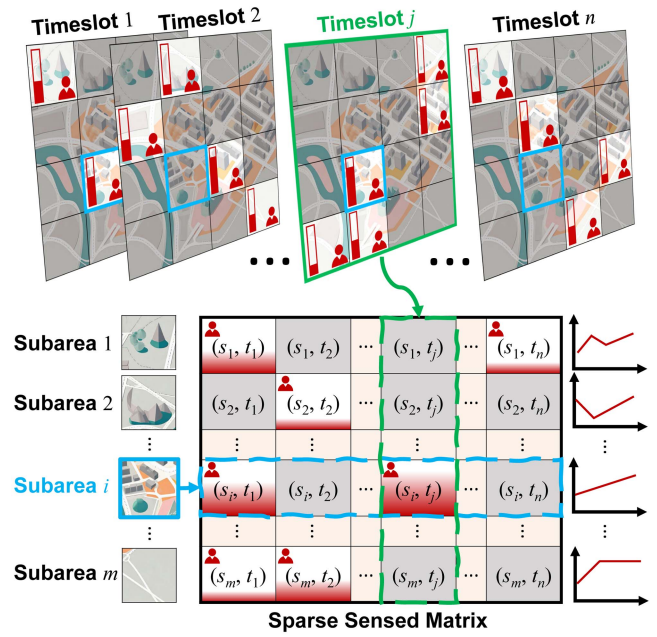


Fig. 3. The formal correspondence between the physical sensing process and mathematical abstraction in canonical crowdsensing paradigms.

focuses on knowledge transfer across similar distributions, exemplified by Rezaee et al.'s [21] UAV-based social distance monitoring using lightweight ShuffleNet. Current transfer learning or domain generalization approaches are not designed for the matrix completion task under extreme sparsity. Consequently, one-shot data completion methods adapted from them lack dynamic model adaptation mechanisms for evolving timeslots under cold-start conditions. Existing methodologies exhibit significant limitations when confronted with novel cold-start tasks.

III. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

The fundamental challenge in sparse crowdsensing systems lies in achieving robust one-shot data completion under extreme cold-start conditions. Traditional approaches falter when confronted with severely limited initial measurements and unpredictable spatiotemporal relationships across disjoint sensing areas. To overcome these constraints, this research adopts a principled mathematical formalization anchored in physical sensing dynamics. As conceptualized in Fig. 1, this research employs a canonical crowdsensing paradigm mobilizing mobile device users to acquire constrained subarea measurements for one-shot data completion (i.e., operates under one-shot conditions, where inference begins with a small set of initial target data (e.g., $n = 1$ timeslot) but leverages source correlations to reduce overall data needs). To realize this objective, we formalize a mathematical framework detailed in this section. Fig. 3 establishes formal correspondence between physical sensing processes and mathematical representations, while Table I enumerates all critical notational conventions adopted throughout this paper.

The objective is to reconstruct complete urban sensing maps comprising m subareas using sparse observations from $\tilde{m}$ sensed

TABLE I
MAIN NOTATIONS

Symbol	Type	Meaning
m, n	$\mathbb{N}_+$	# subareas / timeslots in T.D.
p, q	$\mathbb{N}_+$	# subareas / timeslots in S.D.
i, j	$\mathbb{N}_+$	index of a subarea / timeslot
r	$\mathbb{N}_+$	assumed rank of the matrix
$\mathbf{Y}'_T, \mathbf{Y}_T$	$\mathbb{R}^{m \times n}$	sensed sparse data and ground truth of the complete data in T.D.
$\mathbf{C}_T$	$\{0, 1\}^{m \times n}$	logical label matrix that present the sensed grids at every timeslot in T.D.
$\mathbf{Y}'_S, \mathbf{Y}_S$	$\mathbb{R}^{p \times q}$	sensed sparse data and ground truth of the complete data in S.D.
$\mathbf{C}_S$	$\{0, 1\}^{p \times q}$	logical label matrix that present the sensed grids at every timeslot in S.D.
$\mathbf{Z}_{T1}, \mathbf{Z}_{S1}$	$\mathbb{R}^{r \times (q-n)}$	latent representation for pretrained models by data in T.D. / S.D.
$\mathbf{Z}_{T2}, \mathbf{Z}_{S2}$	$\mathbb{R}^{r \times n}$	latent representation for dedicated / generalized model by data in T.D. / S.D.
$\tilde{\mathbf{Z}}_T$	$\mathbb{R}^{r \times n}$	latent representation for generalized model by data in T.D.
σ_D, σ_G	$\mathbb{R}_+$	RMSE of verification dataset for dedicated / generalized model by data in T.D.
σ_F	$\mathbb{R}_+$	RMSE of data fusion result
$P_T(\cdot), P_S(\cdot)$	function	mathematical representation of pretrained model by data in T.D. / S.D.
$D(\cdot), G(\cdot)$	function	mathematical representation of dedicated / generalized model
$F(\cdot)$	function	mathematical representation of data fusion model
$\hat{\mathbf{Y}}_T$	$\mathbb{R}^{m \times n}$	final estimated data for T.D.

denotes “the number of ...”.

T.D. and **S.D.** represent the target dataset and source dataset, respectively.

subareas ($\tilde{m} \ll m$) per timeslot. Each subarea i is represented by a canonical basis vector $\mathbf{e}^{(i)} \in \{0, 1\}^{m \times 1}$ (one-hot encoding), where only the i^{th} entry is 1. The sensing state is encoded by a binary indicator $c_{ij} \in \{0, 1\}$, with $c_{ij} = 1$ if subarea i is sensed at timeslot j (yielding observation y'_{ij}), and $c_{ij} = 0$ otherwise (where $y'_{ij} = \infty$ denotes missing data). This formalism generates: Sensing state vector $\mathbf{c}^{(j)} = \sum_{i=1}^m c_{ij} \mathbf{e}^{(i)} \in \{0, 1\}^{m \times 1}$; Sparse observation vector $\mathbf{y}'^{(j)} = \sum_{i=1}^m y'_{ij} \mathbf{e}^{(i)} \in \mathbb{R}^{m \times 1}$; Ground truth vector $\mathbf{y}^{(j)} = [y_{1j}, \dots, y_{mj}]^T \in \mathbb{R}^{m \times 1}$. Aggregating n timeslots yields three fundamental matrices: Sensing state matrix $\mathbf{C} = [\mathbf{c}^{(1)} | \dots | \mathbf{c}^{(n)}] \in \{0, 1\}^{m \times n}$; Sparse observation matrix $\mathbf{Y}' = [\mathbf{y}'^{(1)} | \dots | \mathbf{y}'^{(n)}] \in \mathbb{R}^{m \times n}$; Ground truth matrix $\mathbf{Y} = [\mathbf{y}^{(1)} | \dots | \mathbf{y}^{(n)}] \in \mathbb{R}^{m \times n}$. These matrices are intrinsically related through the equation:

$$\mathbf{Y}' = \mathbf{Y} \circ \mathbf{C}, \quad (1)$$

where $\circ$ denotes the Hadamard product between conformable matrices. The primary objective is to establish a data inference function $\Phi(\cdot) : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$ capable of reconstructing unobserved entries from the sparse observation matrix $\mathbf{Y}'$. This yields an estimated complete matrix $\hat{\mathbf{Y}}$ satisfying:

$$\hat{\mathbf{Y}} = \Phi(\mathbf{Y}'). \quad (2)$$

Our methodology transcends exclusive reliance on the target dataset by integrating auxiliary source data to enhance sensing fidelity. While using complete source datasets yields optimal performance, we maximize information extraction from sparse

source data when available. To distinguish data domains, we introduce notational conventions: Target dataset matrices $\mathbf{C}_T$, $\mathbf{Y}_T$, $\mathbf{Y}'_T$, and $\hat{\mathbf{Y}}_T$; Source dataset matrices: $\mathbf{C}_S$, $\mathbf{Y}_S$, $\mathbf{Y}'_S$, and $\hat{\mathbf{Y}}_S$.

Critically, the target and source datasets possess disjoint dimensional profiles. We define that $m \times n$ is the shape of target matrices, and the shape of source matrices is regarded as $p \times q$. i.e.,

$$\begin{cases} \mathbf{C}_T \in \{0, 1\}^{m \times n}, \\ \mathbf{Y}_T, \mathbf{Y}'_T, \hat{\mathbf{Y}}_T \in \mathbb{R}^{m \times n}, \\ \mathbf{C}_S \in \{0, 1\}^{p \times q}, \\ \mathbf{Y}_S, \mathbf{Y}'_S, \hat{\mathbf{Y}}_S \in \mathbb{R}^{p \times q}. \end{cases} \quad (3)$$

B. Problem Formulation

Problem [One-Shot Data Completion via Sparse Crowdsensing]: Given a sparse target observation matrix $\mathbf{Y}'_T \in \mathbb{R}^{m \times n}$ with limited timeslots ($n \ll q$) and its binary indicator matrix $\mathbf{C}_T \in \{0, 1\}^{m \times n}$, alongside a sparse source matrix $\mathbf{Y}'_S \in \mathbb{R}^{p \times q}$ with extensive historical records ($q \gg n$) and $\mathbf{C}_S \in \{0, 1\}^{p \times q}$, we pursue dual objectives:

- **Formalizing the Inference Function:** We define a *functional mapping* $\phi(\cdot) : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$ that reconstructs a complete matrix $\hat{\mathbf{Y}}_T$ from sparse inputs: $\hat{\mathbf{Y}}_T = \phi(\mathbf{Y}'_T | \mathbf{C}_T, \mathbf{Y}'_S, \mathbf{C}_S)$, where $\phi(\cdot)$ operates as a *cross-domain inference operator* synthesizing $\mathbf{Y}'_T \in \mathbb{R}^{m \times n}$, $\mathbf{C}_T \in \{0, 1\}^{m \times n}$, $\mathbf{Y}'_S \in \mathbb{R}^{p \times q}$, and $\mathbf{C}_S \in \{0, 1\}^{p \times q}$.
- **Precision Guarantee Constraint:** The reconstruction must satisfy *bounded-error characteristics* ensuring: $\|\hat{\mathbf{Y}}_T \circ \mathbf{C}_T - \mathbf{Y}'_T\|_F \leq \delta$, with $\delta \rightarrow 0$, guaranteeing *spatiotemporal consistency* and *statistical reliability*.

The core optimization objective entails minimizing the target functional $\mathcal{J}_{\text{obj}}$ defined as:

$$\mathcal{J}_{\text{obj}} = \varepsilon(\phi(\mathbf{Y}'_T | \mathbf{C}_T, \mathbf{Y}'_S, \mathbf{C}_S), \mathbf{Y}_T), \quad (4)$$

where $\varepsilon(\cdot, \cdot)$ quantifies the unsensed entry estimation error, serving as the accuracy metric for sparse completion. Here, $\mathbf{Y}_T$ denotes the ground-truth complete target matrix. It is important to note that $\mathbf{Y}_T$ is used strictly to formalize the recovery target and for subsequent experimental evaluation. It is not an accessible input to the proposed algorithm, which operates solely on the observable data: $\mathbf{Y}'_T$, $\mathbf{C}_T$, $\mathbf{Y}'_S$, and $\mathbf{C}_S$. To achieve $\min_{\phi} \mathcal{J}_{\text{obj}}$, we implement a two-phase optimization paradigm: the **source-domain error minimization** step and the **target-domain fine-tuning** step.

This work extends the few-shot problem formulated in [25] to a one-shot setting, where the goal is to achieve spatiotemporal completion with no target-domain training data. The proposed MapT-STC framework addresses this by integrating cross-domain mappings and dynamic fusion, as detailed in Sections IV and V.

C. The Role of Inter-Data Correlations in MapT-STC

The core innovation of MapT-STC lies in its explicit modeling and utilization of inter-data correlations between the source and target domains. These correlations are not assumed to be static

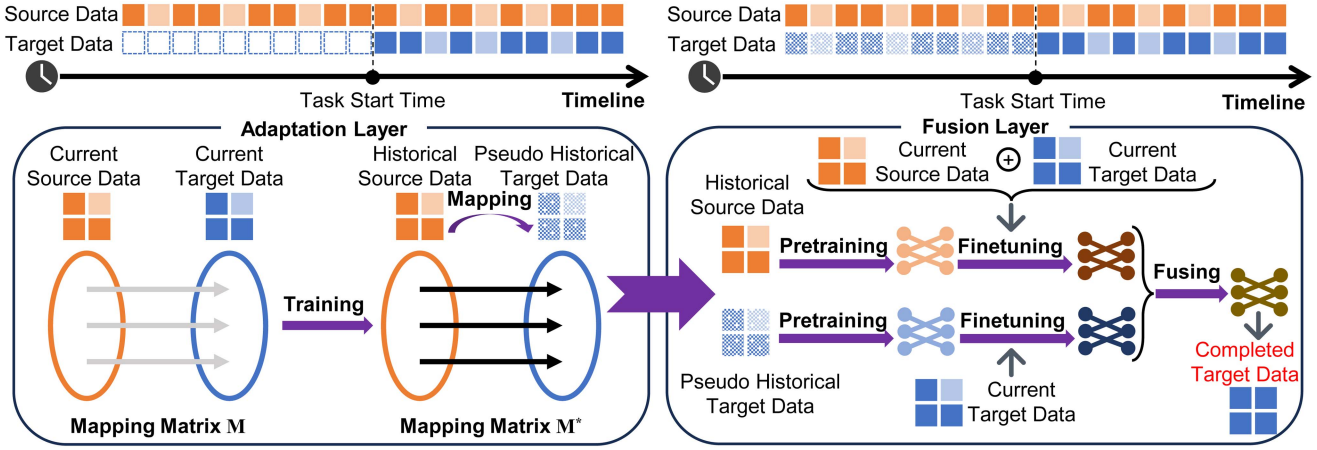


Fig. 4. Framework overview.

(e.g., a fixed Pearson coefficient) but are dynamically captured through domain adaptation. Fig. 4 illustrates the pipeline where inter-data correlations are first captured by the CORAL module and then utilized throughout the dual-model training and fusion process.

We employ the CORAL algorithm to align the feature distributions of the source dataset $\mathbf{Y}'_S$ and the nascent target dataset $\mathbf{Y}'_T$. This alignment, generating pseudo-historical target data, is our method's way of quantifying and leveraging the latent correlation between the two domains. A successful alignment implies a strong underlying correlation that can be transferred.

The aligned features guide the training of the dual-path model. The generalized model $G(\cdot)$ learns patterns that are robust across the correlated domains, while the dedicated model $D(\cdot)$ specializes in the target domain. The strength of the correlation directly influences the reliability of the generalized model's predictions.

The Kalman-filter-based fusion mechanism dynamically weights the contributions of $G(\cdot)$ and $D(\cdot)$. The uncertainty estimates (σ_G^2, σ_D^2) reflect, in part, the persistence and usefulness of the inferred correlations over time. A strong, stable correlation keeps the weight on the generalized model higher for longer.

The effectiveness of the aforementioned correlation-driven mechanisms is contingent upon the judicious selection of a source dataset that exhibits a learnable relationship with the target domain. Therefore, we posit the following operational criteria for identifying a suitable source dataset $\mathbf{Y}'_S$ for a given target $\mathbf{Y}'_T$:

- The source and target should measure phenomena within a shared semantic context (e.g., urban environmental parameters), ensuring that any correlation discovered by CORAL is physically meaningful and not spurious. This foundational relevance is a prerequisite for positive transfer.
- The source dataset must possess sufficient volume and diversity ($q \gg n$) to enable the CORAL algorithm to robustly estimate and align the feature distributions. A source with extensive historical data provides a richer basis for the model to isolate invariant patterns and discard domain-specific noise.

- While we do not pre-suppose a specific correlation strength or form (e.g., linear Pearson correlation), the domains should not be fundamentally disjoint. The source should offer a feature space that, when aligned, can meaningfully constrain the solution space for the target task, as illustrated by the conceptual mappings in Fig. 2. This ensures that the pseudo-historical data generated via $M^*(\cdot)$ provides a productive initialization for the dedicated model.

IV. PRETRAINED MODELS BY THE SOURCE AND THE TARGET DATASET

To address the fundamental challenges outlined in the Introduction, particularly Challenge (i) pertaining to efficient cross-task knowledge transfer under one-shot conditions, we propose a novel framework (Fig. 4) implementing a triple-phase optimization paradigm: First, we construct a generalized completion model by pretraining on sparse source data $\mathbf{Y}'_S$ to capture domain-invariant spatiotemporal patterns. This provides an optimized initialization for cold-start scenarios. Second, we perform simultaneous cross-domain adaptation by: (i) fine-tuning this model with limited target data $\mathbf{Y}'_T$, and (ii) training a dedicated target-specific model to capture local dynamics. Finally, a Kalman Filter-based fusion mechanism dynamically balances the outputs of both models using real-time uncertainty estimates, thereby minimizing the reconstruction error throughout deployment.

Diverging from conventional matrix completion objectives, our primary aim is to derive not the complete source matrix itself, but its latent model representation—encapsulated in the function $P_S(\cdot)$ through training on source data. This learned representation is then transferred to initialize and refine the generalized completion function $G(\cdot)$. The model $P_T(\cdot)$ is the pretrained model on the target domain. It is trained not on real historical target data (which is unavailable in cold-start), but on the pseudo-historical target data generated by the CORAL alignment module $M^*(\cdot)$ (described in Algorithm 1, line 1). This process allows $P_T(\cdot)$ to capture initial patterns specific

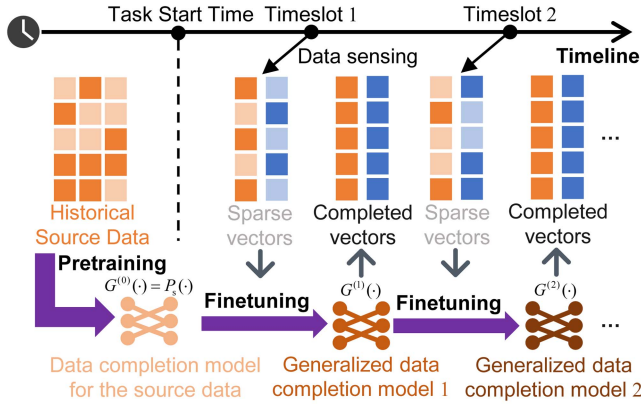


Fig. 5. The structure and training process of the generalized model.

to the target domain's feature distribution, as transferred from the source domain via $M^*(\cdot)$. It serves as the foundation for initializing the dedicated model $D(\cdot)$ later. As depicted in Fig. 5, we innovate upon Deep Matrix Factorization (DMF) [18] to architect a novel transfer learning framework for sparse matrix completion. The algorithmic implementation of this pretrained model is elaborated in subsequent sections, with a comprehensive pseudo code specification provided in Algorithm 1 for enhanced reproducibility.

The core idea of Algorithm 1 is to bootstrap the cold-start learning process by constructing two pretrained models: one ($P_S(\cdot)$) learns general spatiotemporal patterns from the abundant source data, while the other ($P_T(\cdot)$) is trained on pseudo-historical target data generated by the CORAL module $M^*(\cdot)$. This generation aligns the feature distribution of the source domain with the very limited target observations (e.g., a single timeslot), effectively creating a synthetic history for the target task. These pretrained models, $P_S(\cdot)$ and $P_T(\cdot)$, serve as optimized initializations for the generalized model $G(\cdot)$ and the dedicated model $D(\cdot)$, respectively. This two-model initialization strategy is the cornerstone of our one-shot framework, enabling rapid adaptation by transferring knowledge from the correlated source domain while building a specialized model for the target.

The pretrained models $P_S(\cdot)$ and $P_T(\cdot)$ are built upon the architecture of DMF [18]. While standard DMF is designed for classical sparse crowdsensing tasks and is not directly applicable to the one-shot completion problem, its model structure is well-suited for learning latent representations from data. Therefore, we adapt the DMF framework to construct our pretraining modules. Crucially, we leverage the trained network structure and parameters of DMF, rather than its direct data completion outputs, to initialize $P_S(\cdot)$ and $P_T(\cdot)$. This approach allows us to transfer the representational capacity learned from the source domain and the aligned pseudo-target domain, forming the foundation for the subsequent dual-path learning process detailed in Algorithm 1.

The initial phase entails structuring observational data into subarea-timeslot matrices $\mathbf{Y}'_S \in \mathbb{R}^{p \times q}$ and $\mathbf{Y}'_T \in \mathbb{R}^{m \times n}$, where rows encode subarea indices, columns represent timeslots, and elements store observational metrics. The corresponding binary indicator matrices $\mathbf{C}_S \in \mathbb{R}^{p \times q}$ and $\mathbf{C}_T \in \mathbb{R}^{m \times n}$ identify

Algorithm 1: Pretrained Models by the Source and the Target Dataset.

Input: $\mathbf{Y}'_S = [\mathbf{y}'_S^{(-q+n)} | \dots | \mathbf{y}'_S^{(-1)} | \mathbf{y}'_S^{(1)} | \dots | \mathbf{y}'_S^{(n)}]$,

$\mathbf{C}_S = [\mathbf{c}_S^{(-q+n)} | \dots | \mathbf{c}_S^{(-1)} | \mathbf{c}_S^{(1)} | \dots | \mathbf{c}_S^{(n)}]$,

$\mathbf{Y}'_T = [\mathbf{y}'_T^{(1)} | \dots | \mathbf{y}'_T^{(n)}]$, $\mathbf{C}_T = [\mathbf{c}_T^{(1)} | \dots | \mathbf{c}_T^{(n)}]$.

Output: the model parameters $P_S(\cdot)$ and $P_T(\cdot)$.

* Preparing *\

- 1: Train the Neural Network $M(\cdot)$ with $[\mathbf{y}'_S^{(-q+n+1)} | \dots | \mathbf{y}'_S^{(q)}]$ as input and $[\mathbf{y}'_T^{(1)} | \dots | \mathbf{y}'_T^{(n)}]$ as output, then obtain $M^*(\cdot)$;
 - 2: Random Init. $\mathbf{Z}_{S1} = [\mathbf{z}_S^{(-q+n)} | \dots | \mathbf{z}_S^{(-1)}]$ and the NN structure parameters of $P_S(\cdot)$;
 - 3: Random Init. $\mathbf{Z}_{T1} = [\mathbf{z}_T^{(-q+n)} | \dots | \mathbf{z}_T^{(-1)}]$ and the NN structure parameters of $P_T(\cdot)$;
 - * Training *\
 - 4: $count \leftarrow 0$;
 - 5: **while** not convergent **and** $count < \text{MAX_ITER}$ **do**
 - 6: **for** j is from $(-q+n)$ to -1 **do**
 - 7: Calculate $\|P_S(\mathbf{z}_S^{(j)}) \circ \mathbf{c}_S^{(j)} - \mathbf{y}'_S^{(j)}\|$;
 - 8: Fix $\mathbf{z}_S^{(j)}$ and the NN structure parameters of $P_S(\cdot)$ at the same time to make $\mathcal{L}_{PS}^{(j)}$ reduce;
 - 9: Calculate $\|P_T(\mathbf{z}_T^{(j)}) \circ \mathbf{c}_S^{(j)} - M^*(\mathbf{y}'_S^{(j)})\|$;
 - 10: Fix $\mathbf{z}_T^{(j)}$ and the NN structure parameters of $P_T(\cdot)$ at the same time to make $\mathcal{L}_{PT}^{(j)}$ reduce;
 - 11: **end for**
 - 12: $count \leftarrow count + 1$;
 - 13: **end while**
 - * Outputting *\
 - 14: **return** $P_S(\cdot)$ and $P_T(\cdot)$.
-

valid entries. In order to create pseudo historical data (index from $(-q+n)$ to -1) for the target domain, we introduce the correlation alignment (CORAL) module $M(\cdot)$ with $[\mathbf{y}'_S^{(-q+n+1)} | \dots | \mathbf{y}'_S^{(q)}]$ as input and $[\mathbf{y}'_T^{(1)} | \dots | \mathbf{y}'_T^{(n)}]$ as output. This module is designed to tackle Challenge (i) by learning to align feature distributions between domains, thereby enabling knowledge transfer from the source to the target with minimal target data. Through cold-start phase data (index from 1 to n), $M(\cdot)$ can be updated to $M^*(\cdot)$. Then, we implement deep factorization architectures $P_S(\cdot)$ and $P_T(\cdot)$ comprising multiple non-linear layers to learn latent representations. This neural network maps the input matrix to a lower-dimensional embedding space through hierarchical feature abstraction $\mathbf{Z}_S = [\mathbf{Z}_{S1} | \mathbf{Z}_{S2}]$ and $\mathbf{Z}_T = [\mathbf{Z}_{T1} | \mathbf{Z}_{T2}]$. For the details of these matrices:

$$\begin{cases} \mathbf{Z}_{S1} = [\mathbf{z}_S^{(-q+n)} | \dots | \mathbf{z}_S^{(-1)}], \mathbf{Z}_{S2} = [\mathbf{z}_S^{(1)} | \dots | \mathbf{z}_S^{(n)}], \\ \mathbf{Z}_{T1} = [\mathbf{z}_T^{(-q+n)} | \dots | \mathbf{z}_T^{(-1)}], \mathbf{Z}_{T2} = [\mathbf{z}_T^{(1)} | \dots | \mathbf{z}_T^{(n)}]. \end{cases} \quad (5)$$

Prior to factorization, latent factors require initialization of $P_S(\cdot)/P_T(\cdot)$ and $\mathbf{Z}_{S1}/\mathbf{Z}_{T1}$. Following established deep learning practice, these parameters are typically initialized via Gaussian distribution sampling. The training procedure minimizes the

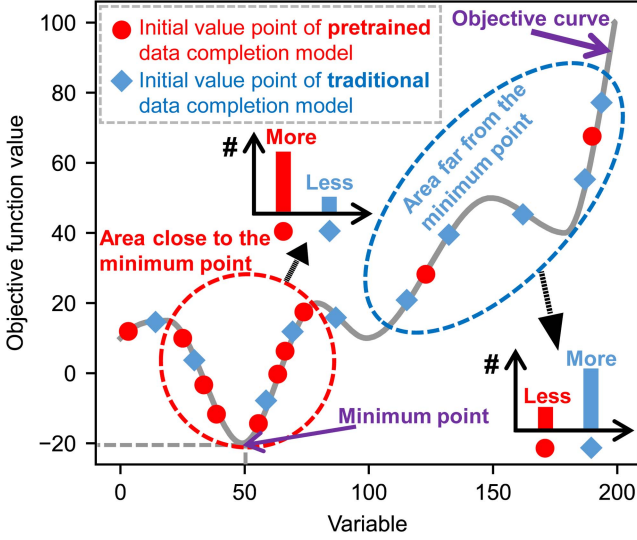


Fig. 6. The divergence in initialization configurations between the two data completion models' parameter spaces.

reconstruction loss $\mathcal{L}_{PS}^{(j)}/\mathcal{L}_{PT}^{(j)}$ at each timeslot j through:

$$\mathcal{L}_{PS}^{(j)} = \|P(\mathbf{z}_S^{(j)}) \circ \mathbf{c}_S^{(j)} - \mathbf{y}_S'^{(j)}\|_F^2 + \lambda_1 \pi(P_S) + \lambda_2 \|\mathbf{z}_S^{(j)}\|_F^2, \quad (6)$$

$$\mathcal{L}_{PT}^{(j)} = \|P_T(\mathbf{z}_T^{(j)}) \circ \mathbf{c}_S^{(j)} - M^*(\mathbf{y}_S'^{(j)})\|_F^2 + \lambda_3 \pi(P_T) + \lambda_4 \|\mathbf{z}_T^{(j)}\|_F^2, \quad (7)$$

where $\|\cdot\|_F$ denotes the Frobenius Norm of a matrix. $\pi(\cdot)$ is a functional function used to represent penalty terms related to network parameters. $\lambda_1, \dots, \lambda_4$ denote the weight coefficients of the model and the inputs, respectively.

Gradient updates derived from the loss function propagate backward through the neural network via backpropagation. Critically, (6) and (7) reveal that the latent vector $\mathbf{z}_S^{(j)}/\mathbf{z}_T^{(j)}$ is embedded in both the optimization objective and the regularization component. This configuration enables the network's weights to be tuned for optimal latent factor extraction from the input matrix. These extracted latent factors—represented by the network's output—serve as the foundation for decomposing the original matrix into two reduced-dimensionality matrices. The model predicts missing entries by reconstructing the matrix through the multiplicative recombination of these factor matrices.

In machine learning, optimizing the target function is the primary task to achieve a high-performing model. Our target function curve (as shown in Fig. 6) has a global minimum that needs to be found to design an optimal model. To find this global minimum point, we generally use uniform random sampling as initial values for the parameters of our model, followed by gradient descent optimization. However, with the help of source datasets, we can use pretrained models' parameters as initial values for training the target dataset. This parameter initialization strategy facilitates convergence near the global minimum, lowering initial data completion errors. In this article, we will analyze this approach using relevant methods and techniques.

Leveraging source datasets for model pretraining is a widespread practice that enhances model performance and curtails training duration. The pretraining method involves training our model using a source dataset, which is relevant to our target dataset's problem or task. The idea is to train the model using the source dataset and then fine-tune it using our target dataset. This process allows the model to learn from the insights gained from the source dataset, which can be used to improve its performance on our target dataset. As established, pretraining enables the model to assimilate knowledge from the source dataset, yielding an effective initialization of parameter values within near-optimal ranges. Since parameters derived from pretraining reside near the global minimum of the source dataset's loss landscape, they exhibit higher probability of proximity to the global minimum for the target dataset. This approach mitigates the risk of initializing parameters distant from global optima, consequently reducing initial-phase error.

V. GENERALIZED MODEL, DEDICATED MODEL AND THEIR FUSION MODEL FOR MATRIX COMPLETION

Following the pretraining methodology in Section IV, we obtain source/target data completion models $P_S(\cdot)/P_T(\cdot)$. These pretrained models serve as initializations for training the generalized model $G(\cdot)$ and dedicated model $D(\cdot)$. This section details the methodology for training $G(\cdot)$ (using both source and target datasets) and $D(\cdot)$ (using only the target dataset). The dual-path learning system (comprising $G(\cdot)$ and $D(\cdot)$) is central to addressing Challenge (ii), as it dynamically establishes an equilibrium between similarity preservation (via the generalized model) and diversity adaptation (via the dedicated model) to optimize domain transfer and reduce temporal dependencies. The model architecture is illustrated in Fig. 7, with algorithmic implementation provided in Algorithm 2.

This section's methodology employs dual sub-modules to resolve data sparsity and incompleteness within the target dataset. The model processes temporal sequence fragments from the target dataset, coordinating simultaneous optimization of both sub-components. The primary sub-module trains a data completion module capable of handling missing data in both target and source datasets. Leveraging the pretrained model $P_S(\cdot)$ from historical source data, we initialize the generalized module $G(\cdot)$. As $G(\cdot)$ ingests temporal sequences from either dataset, its weights and parameters adapt through classical DMF techniques – which analyze input patterns, decode trends, and impute missing values – progressively evolving into a generalized data completion module. Consequently, the generalized model's loss function is defined as:

$$\begin{aligned} \mathcal{L}_G^{(j)} = & \|G(\tilde{\mathbf{z}}_T^{(j)}) \circ \mathbf{c}_T^{(j)} - \mathbf{y}_T'^{(j)}\|_F^2 + \lambda_5 \|G(\mathbf{z}_S^{(j)}) \circ \mathbf{c}_S^{(j)} - \mathbf{y}_S'^{(j)}\|_F^2 \\ & + \lambda_6 \pi(G) + \lambda_7 \|\tilde{\mathbf{z}}_T^{(j)}\|_F^2 + \lambda_8 \|\mathbf{z}_S^{(j)}\|_F^2, \end{aligned} \quad (8)$$

where $\lambda_5, \dots, \lambda_8$ denote the weight coefficients.

The secondary sub-module constitutes a dedicated model trained exclusively on the target dataset. Deploying DMF techniques, this component constructs localized embeddings to impute missing observations within the target data space. Its loss

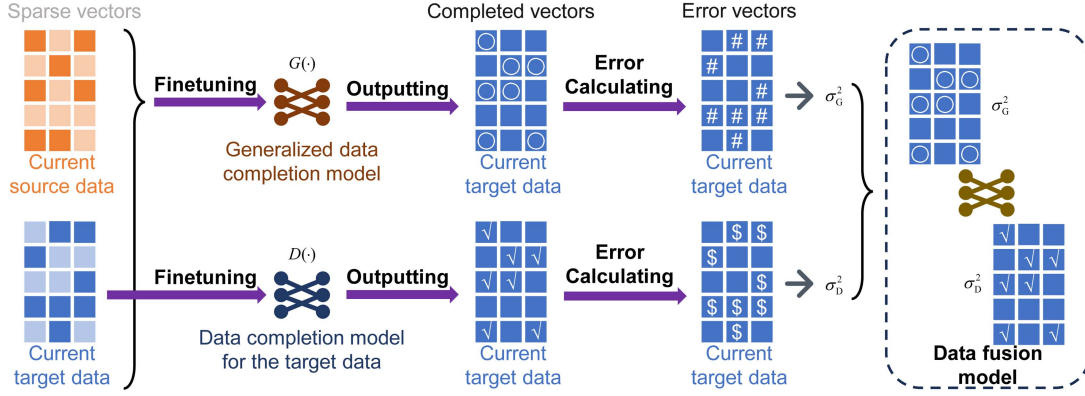


Fig. 7. Architectural design of the target data completion and data fusion modules.

Algorithm 2: Generalized Model, Dedicated Model and Their Fusion Model for Matrix Completion.

Input: $\mathbf{Y}'_S = [\mathbf{y}'^{(-q+n)}_S | \dots | \mathbf{y}'^{(-1)}_S | \mathbf{y}'^{(1)}_S | \dots | \mathbf{y}'^{(n)}_S]$,
 $\mathbf{C}_S = [\mathbf{c}_S^{(-q+n)} | \dots | \mathbf{c}_S^{(-1)} | \mathbf{c}_S^{(1)} | \dots | \mathbf{c}_S^{(n)}]$,
 $\mathbf{Y}'_T = [\mathbf{y}'^{(1)}_T | \dots | \mathbf{y}'^{(n)}_T]$, $\mathbf{C}_T = [\mathbf{c}_T^{(1)} | \dots | \mathbf{c}_T^{(n)}]$, and the
 pretrained model $P_S(\cdot)$ and $P_T(\cdot)$.
Output: the completed matrix $\hat{\mathbf{Y}}_T$.

```

  \* Preparing *
1: Random Init. input vectors  $\mathbf{Z}_{S2} = [\mathbf{z}_S^{(1)} | \dots | \mathbf{z}_S^{(n)}]$ ,  

    $\mathbf{Z}_{T2} = [\mathbf{z}_T^{(1)} | \dots | \mathbf{z}_T^{(n)}]$ ,  $\tilde{\mathbf{Z}}_T = [\tilde{\mathbf{z}}_T^{(1)} | \dots | \tilde{\mathbf{z}}_T^{(n)}]$ ;
2:  $G(\cdot) \leftarrow P_S(\cdot)$ ,  $D(\cdot) \leftarrow P_T(\cdot)$ ;
   \* Training *
3:  $count \leftarrow 0$ ;
4: while not convergent and  $count < \text{MAX\_ITER}$  do
5:   for  $j$  is from 1 to  $n$  do
6:     Calculate  $\|G(\mathbf{z}_S^{(j)}) \circ \mathbf{c}_S^{(j)} - \mathbf{y}'^{(j)}_S\|$ ;
7:     Calculate  $\|G(\tilde{\mathbf{z}}_T^{(j)}) \circ \mathbf{c}_T^{(j)} - \mathbf{y}'^{(j)}_T\|$ ;
8:     Fix  $\mathbf{z}_S^{(j)}$ ,  $\tilde{\mathbf{z}}_T^{(j)}$  and the NN structure parameters of  

        $G(\cdot)$  at the same time to make  $\mathcal{L}_G^{(j)}$  reduce;
9:     Calculate  $\|D(\mathbf{z}_T^{(j)}) \circ \mathbf{c}_T^{(j)} - \mathbf{y}'^{(j)}_T\|$ ;
10:    Fix  $\mathbf{z}_T^{(j)}$  and the NN structure parameters of  $D(\cdot)$   

       at the same time to make  $\mathcal{L}_D^{(j)}$  reduce;
11:   end for
12:   Calculate  $\sigma_G$  and  $\sigma_D$  by (10) and (11), respectively;
13:   Calculate  $\hat{\mathbf{Y}}_T$  by (13);
14:    $count \leftarrow count + 1$ ;
15: end while
   \* Outputting *
16: return  $\hat{\mathbf{Y}}_T$ .
```

function is correspondingly defined as:

$$\mathcal{L}_D^{(j)} = \|D(\mathbf{z}_T^{(j)}) \circ \mathbf{c}_T^{(j)} - \mathbf{y}'^{(j)}_T\|_F^2 + \lambda_9 \pi(D) + \lambda_{10} \|\mathbf{z}_T^{(j)}\|_F^2, \quad (9)$$

where λ_9 and λ_{10} denote the weight coefficients. Unlike the primary sub-module, this component eschews generalization to external datasets, instead concentrating exclusively on optimizing spatiotemporal data completion.

The outputs from both sub-modules undergo fusion via a data fusion module, which employs Kalman Filtering techniques to consolidate their predictions into a unified matrix representation. This adaptive, uncertainty-aware fusion mechanism directly addresses Challenge (iii) by constructing dynamic, data-driven couplings between models to replace static correlation metrics, thereby minimizing fine-tuning costs in cold-start scenarios. This matrix accurately reconstructs the original dataset's latent trends and patterns, overcoming inherent data incompleteness and sparsity. Crucially, the Kalman Filter dynamically computes the completion error for each sub-module and allocates their contributions via covariance-driven weighting, optimizing fusion fidelity. The Mean Squared Error (MSE) for this process is calculated as follows:

$$\sigma_G^2 = \frac{1}{|\mathcal{A}|} \|G(\tilde{\mathbf{Z}}_T) \circ \mathbf{C}_T - \mathbf{Y}'_T\|_F^2, \quad (10)$$

$$\sigma_D^2 = \frac{1}{|\mathcal{A}|} \|D(\mathbf{Z}_T) \circ \mathbf{C}_T - \mathbf{Y}'_T\|_F^2, \quad (11)$$

where $\mathcal{A} = \{(i, j) | c_{ij} = 1, 1 \leq i \leq m, 1 \leq j \leq n, (i, j) \in \mathbb{N}_+^2\}$ while $|\mathcal{A}|$ denotes the number of elements in the set $\mathcal{A}$. Through Kalman Filter-based data fusion, the final fused data error σ_F conforms to the following mathematical relation:

$$\frac{1}{\sigma_F^2} = \frac{1}{\sigma_D^2} + \frac{1}{\sigma_G^2}. \quad (12)$$

It is obvious that the fusion data error σ_{fus} is smaller than both σ_D and σ_G . This is consistent with the theoretical analysis results in Fig. 8. As shown in Fig. 8, the maximum point of the probability density curve after data fusion is between two sub-modules, and the peak value is higher than both the dedicated module and the generalized module.

Finally, we will get the final result of data fusion through the following matrix formula:

$$\hat{\mathbf{Y}}_T = \frac{\sigma_D^2}{\sigma_G^2 + \sigma_D^2} G(\tilde{\mathbf{Z}}_T) + \frac{\sigma_G^2}{\sigma_G^2 + \sigma_D^2} D(\mathbf{Z}_T). \quad (13)$$

Following the initial mapping and alignment via CORAL, the framework enables a chain-style propagation of inferences

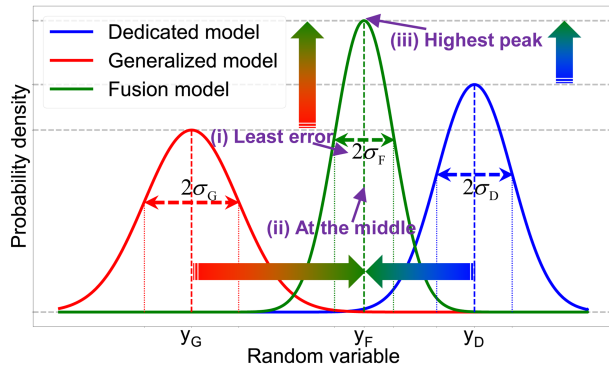


Fig. 8. Probability distribution coupling in data fusion.

across subsequent timeslots. This propagation mechanism leverages the transferred knowledge from the source-domain pre-training and the continuously adapted dual-path models to sustain accurate completion with only the one-shot initial target sample. Thus, the model not only addresses the cold-start challenge but also ensures robust, sequential data recovery in an ongoing manner.

VI. PERFORMANCE EVALUATION

This section systematically characterizes two benchmark datasets for sparse crowdsensing inference tasks, followed by rigorous technical specifications of established and state-of-the-art baseline methodologies. We subsequently benchmark our proposed framework's performance in addressing sparse crowdsensing challenges. To contextualize experimental significance, we pose critical research inquiries that quantitatively assess the method's effectiveness, efficiency, and robustness against the SOTA baseline methods.

- *RQ1*: What performance gains does MapT-STC achieve for the target sparse crowdsensing problems?
- *RQ2*: How does MapT-STC's convergence dynamics evolve during pre-asymptotic training phases?
- *RQ3*: How do the two sub-modules' contributions to MapT-STC change as data volume increases?
- *RQ4*: How does MapT-STC's robustness degrade under varying sensed data ratios?
- *RQ5*: Is every component of MapT-STC necessary?

Through our experiments, we aim to investigate whether our approach can effectively leverage the available data to improve the accuracy of sparse crowdsensing data curation. Moreover, we aim to test the scalability of our proposed approach, particularly in large-scale sparse datasets. Additionally, we seek to assess the extent to which our approach can ameliorate the limitations of existing baseline approaches, particularly in noisy and heterogeneous datasets. Finally, we examine the performance of our approach relative to the state-of-the-art solutions in the domain of data inference for sparse crowdsensing tasks. Overall, this section will offer a detailed account of our experimental setup and methodology, enabling readers to gain a comprehensive understanding of the effectiveness of our proposed approach in addressing the problems and challenges of one-shot data completion in sparse crowdsensing.

TABLE II
STATISTICAL CHARACTERIZATION OF THREE BENCHMARK EVALUATION DATASETS

	Datasets	
	<i>Sensor-Scope</i>	<i>U-Air</i>
Country/Region	Switzerland	Chinese mainland
City	Lausanne	Beijing
Subarea	30m × 50m × 57	1km × 1km × 36
Timeslot	30min × 336	60min × 264
Target data	Temperature	PM _{2.5}
Mean ± Std. [Unit]	6.04 ± 1.87 [°C]	79.11 ± 81.21 [μg/m ³]
Source data	Humidity	PM ₁₀
Mean ± Std. [Unit]	84.52 ± 6.32 [%]	63.12 ± 48.56 [μg/m ³]

A. Datasets

We benchmark MapT-STC on one-shot data completion using two urban crowdsensing datasets: *U-Air* [42] and *Sensor-Scope* [43]. Each dataset encompasses dual correlated tasks, selected for rigorous evaluation of cross-task generalization. Comprehensive statistical characterization is provided in Table II. These datasets are widely used within the research community and have been collected over extended periods of time with high accuracy and precision, making them ideal for evaluating our proposed approach and other baseline methods.

- The *U-Air* [42] dataset is a valuable resource for analyzing air quality in Beijing City, Chinese mainland. It includes important measurements such as PM₁₀ and PM_{2.5}, which are critical factors in understanding air pollution. We selected the PM_{2.5} value to evaluate our one-shot data completion problem, as it is one of the most significant indicators of air quality. To facilitate our analysis, we chose to focus on 36 subareas, each measuring 1 km × 1 km, from the *U-Air* dataset. These subareas were selected based on their relevance to our study and their proximity to our research location. By working with this urban sensing data, we were able to perform comprehensive evaluations of one-shot data completion in the context of air quality analysis. Our methodology was informed by similar studies, such as [42], which provided valuable insights into the challenges of working with urban sensing data. Through our analysis, we hope to contribute to ongoing efforts to improve air quality monitoring and management in urban environments.
- The *Sensor-Scope* [43] dataset includes various types of sensing data collected from static sensors located around the École Polytechnique Fédérale de Lausanne (EPFL) campus, which captures urban environmental conditions. For this study, we focused solely on representative temperature data, although other types of data, such as humidity, were also available. To facilitate statistical analysis of spatially distinct subareas, we divided the EPFL campus into 10 × 10 subareas. Each subarea measures 50m × 30m, resulting in a total of 57 subareas. These subareas capture a range of environmental conditions, allowing for comprehensive data analysis. After performing appropriate data pre-processing, we collected a significant amount of

TABLE III
COMPARATIVE ANALYSIS OF DATA INFERENCE TECHNIQUES

Method	Non-linear	DNN-based	dual-path	one-shot
KNN [15]	×	×	×	×
GP [44], [45]	✓	×	×	×
DMF [18]	✓	✓	×	×
IGMC [20]	✓	✓	×	×
TSJT [46]	✓	✓	✓	×
ImputeFormer [47]	✓	✓	×	×
META-MCS [48]	✓	✓	✓	×
CDFMC [25]	✓	✓	✓	×
MapT-STC *	✓	✓	✓	✓

* It is proposed in Sections IV and V of this paper.

sensing data from these subareas. The aim of our one-shot data completion problem is to mine this data effectively and derive meaningful insights into the impact of urban conditions on environmental factors like temperature or humidity values.

B. Baselines & Measures

1) *Baselines*: To enable one-shot spatiotemporal data inference from sparse sensed data, we propose the MapT-STC algorithm. Its efficacy is benchmarked against diverse data inference methods, including:

- *K-Nearest Neighbor (KNN)* [15] infers values via temporal proximity-based imputation using the data values from the top- K sensed timeslots.
- *Gaussian Process (GP)* [44], [45] constructs spatiotemporal Gaussian distributions to model and generate unsensed observation values.
- *Deep Matrix Factorization (DMF)* [18] leverages neural matrix completion for sparse data reconstruction.
- *Inductive Graph-Based Matrix Completion (IGMC)* [20] employs graph neural networks for relational inference in recommendation systems.
- *Target-Skewed Joint Training (TSJT)* [46] is a cross-city spatiotemporal forecasting model based on transfer learning framework.
- *Low Rankness-Induced Transformers for Generalizable Spatiotemporal Imputation (ImputeFormer)* [47] is a transformer-based architecture utilizing spatiotemporal attention and generative modeling for long-range, multivariate time series imputation.
- *META-MCS* [48] is a meta-knowledge-based multiple data inference framework for sparse crowdsensing.
- *Correlated Data Fusion for Matrix Completion (CDFMC)* [25] employs transfer learning to fine-tune an auxiliary dataset-based pretrained model using minimal objective data.

Understanding the intricate connections and distinct traits among these baselines is evident. As a result, we have compiled a summary of the connections and attributes among these approaches in Table III.

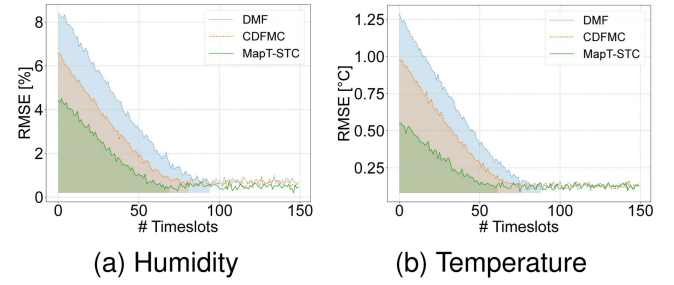


Fig. 9. Inference accuracy under different numbers of timeslots over *Sensor-Scope*.

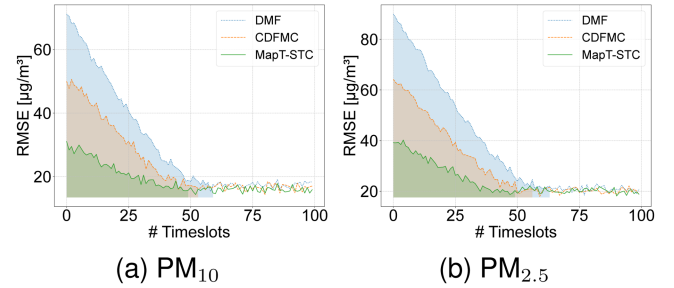


Fig. 10. Inference accuracy under different numbers of timeslots over *U-Air*.

By systematically benchmarking MapT-STC against established baselines, we delineate performance boundaries and inferential limitations across methodologies. This empirical analysis establishes the optimal technical pathway for one-shot spatiotemporal data inference from sparse sensed data, ultimately enhancing engineering viability for data-driven decision systems in mission-critical applications.

2) *Measures*: Our experimental evaluation examines three core metrics:

- *Root Mean Square Error (RMSE)*, which quantifies reconstruction fidelity between inferred and ground-truth values. Lower RMSE signifies higher accuracy.
- *Number of timeslots required*, which measures operational efficiency via minimal sensing epochs needed to achieve error tolerance thresholds. Fewer slots reduce sensing costs.
- *Comprehensive impact*, which denotes the synthesis of impacts from the first two aspects. The geometric meaning is the area between the RMSE functional curve and the timeslot axis.

C. Overall Improvement Analysis (RQ1)

We conducted a cross-dataset multi-task experimentation comparing MapT-STC against the Deep Matrix Factorization (DMF) baseline. Performance was evaluated through RMSE dynamics under varying timeslot allocations across four distinct tasks. Results (Figs. 9 and 10) demonstrate MapT-STC's persistent superiority: it consistently outperformed DMF and CDFMC from initial sensing phases until convergence to the theoretical error bound.

Experimental results confirm MapT-STC's performance enhancement over DMF and CDFMC, evidenced by significant

TABLE IV
 INITIAL ERROR UNDER DIFFERENT METHODS OVER *SENSOR-SCOPE* AND *U-Air*

Method	Task			
	Humi.	Temp.	PM ₁₀	PM _{2.5}
DMF	8.3925	1.2851	71.180	89.974
CDFMC	6.5801	0.9781	49.985	64.164
MapT-STC	4.4056	0.5487	31.072	39.158
Improv. amount	2.1745	0.4294	18.913	25.006
Improv. rate	33.1%	43.9%	37.8%	39.0%

Use **yellow** highlighting to mark the optimal method, and **pink** highlighting to mark the suboptimal method.

 TABLE V
 CONVERGENCE TIME UNDER DIFFERENT METHODS OVER *SENSOR-SCOPE* AND *U-Air*

Method	Task			
	Humi.	Temp.	PM ₁₀	PM _{2.5}
DMF	95	89	60	64
CDFMC	82	77	54	57
MapT-STC	70	62	50	50
Improv. amount	12	15	4	7
Improv. rate	14.6%	19.5%	7.4%	12.3%

Use **yellow** highlighting to mark the optimal method, and **pink** highlighting to mark the suboptimal method.

 TABLE VI
 CUMULATIVE CONTRIBUTION AT CONVERGENCE TIME UNDER DIFFERENT METHODS OVER *SENSOR-SCOPE* AND *U-Air*

Method	Task			
	Humi.	Temp.	PM ₁₀	PM _{2.5}
DMF	358.04	50.8320	2258.1	3038.1
CDFMC	239.42	35.6734	1623.7	2125.6
MapT-STC	147.18	18.3679	1057.9	1365.1
Improv. amount	92.24	17.3055	565.7	760.5
Improv. rate	38.5%	48.5%	34.8%	35.8%

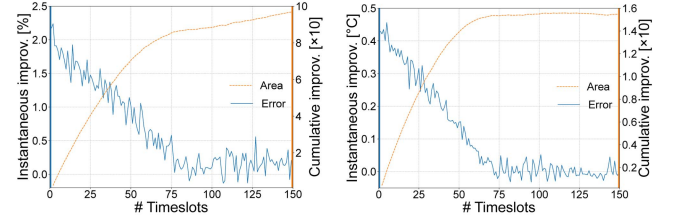
Use **yellow** highlighting to mark the optimal method, and **pink** highlighting to mark the suboptimal method.

RMSE reduction. This improvement stems from MapT-STC's explicit modeling of cross-dataset correlations, which optimizes matrix completion accuracy. Concurrently, both methods exhibit information-gain convergence: increasing timeslot allocations consistently reduce RMSE by expanding the learnable data manifold for both algorithms.

We further benchmark DMF, CDFMC, and MapT-STC along three dimensions: (i) Convergence time (Table V); (ii) Initial error (TABLE IV); (iii) Cumulative contribution (Table VI).

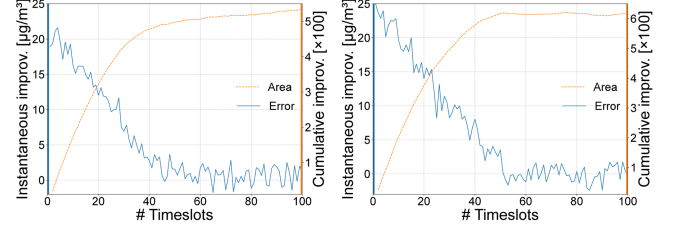
Convergence time is an important factor in determining the efficiency of an algorithm. The results showed that MapT-STC has a superior convergence speed compared to DMF and CDFMC. MapT-STC showed an improvement of over 7% in convergence time. This improvement in convergence speed can be attributed to the advanced data fusion method implemented in MapT-STC. DMF, on the other hand, uses basic matrix factorization that might be slower in the data completion process.

Initial error measures the accuracy of the predicted data compared to the ground truth data. Our results showed that



(a) Humidity

(b) Temperature

 Fig. 11. Instantaneous and cumulative improvement of MapT-STC under different numbers of timeslots over *Sensor-Scope*.

 (a) PM₁₀

 (b) PM_{2.5}

 Fig. 12. Instantaneous and cumulative improvement of MapT-STC under different numbers of timeslots over *U-Air*.

MapT-STC outperformed the suboptimal method CDFMC in terms of initial error. This improvement was significant, with MapT-STC having almost 30% lower initial error than CDFMC. The accuracy of MapT-STC's predicted data is attributed to the advanced pretrained model. Cumulative contribution represents the combined effect of convergence time and initial error. Our results showed a cumulative contribution improvement of about 30% in MapT-STC over CDFMC. This indicates that MapT-STC handled the matrix completion task more efficiently than CDFMC, as it achieved higher accuracy with faster convergence.

The superior results of MapT-STC likely result from its data fusion approach, which leverages information from multiple correlated datasets for matrix completion. This contrasts with the traditional matrix factorization technique used in DMF, leading to more accurate and efficient predictions for MapT-STC. In conclusion, experimental results demonstrate that MapT-STC outperforms DMF regarding Convergence time, Initial Error, and Cumulative Contribution. MapT-STC's enhanced performance can be attributed to the advanced data fusion technique implemented in the algorithm. These findings have significant implications for researchers and practitioners working in the field of matrix completion, as MapT-STC presents a more reliable and efficient alternative to the traditional matrix factorization-based techniques.

D. Macro-Process Analysis (RQ2)

Our comparative analysis of MapT-STC, traditional DMF, and CDFMC for timeslot-based data completion reveals distinct performance patterns. Figs. 11 and 12 demonstrate that MapT-STC exhibits a declining trend in data imputation error (blue solid line), indicating its particular efficacy with limited initial training data. Concurrently, the cumulative contribution curve (yellow

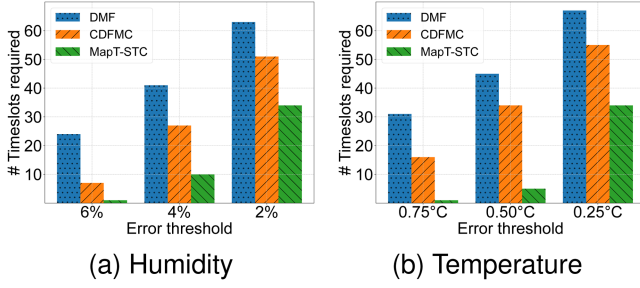


Fig. 13. The number of timeslots required under different error thresholds over *Sensor-Scope*.

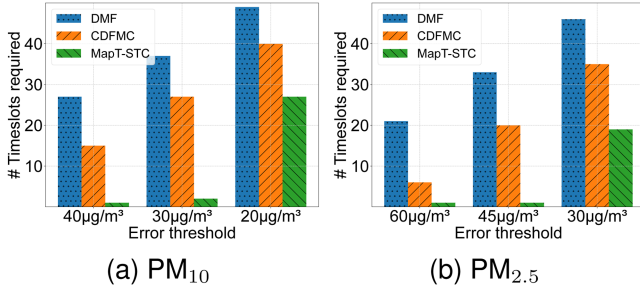


Fig. 14. The number of timeslots required under different error thresholds over *U-Air*.

dotted line) shows an upward trajectory that plateaus post-model convergence, reflecting diminishing returns from MapT-STC thereafter. These findings establish MapT-STC as superior to DMF and CDFMC under training data scarcity, especially valuable for dynamic data adaptation tasks where data uncertainty significantly impacts performance. The cumulative contribution curve further confirms MapT-STC's advantage in capturing data trends pre-convergence. While MapT-STC's benefits attenuate with reduced data sparsity over time, it maintains an accelerated convergence speed crucial for large-scale datasets.

To quantify error tolerance dependencies, we measured the minimum timeslot requirement across varying error thresholds (using DMF as baseline in Figs. 13 and 14). This evaluation framework constitutes a generalizable precision-enhancing paradigm for data completion methodologies, not an endorsement of MapT-STC's superiority. This experiment reveals that for any error threshold of each task, the proposed MapT-STC requires fewer timeslots than the previous DMF and CDFMC. Moreover, as the artificially set error threshold decreases, the number of timeslots required for all methods shows an increasing trend. This is logical and consistent with our intuitive perception. The smaller the error threshold, the stricter the standard, and the greater the amount of data needed to train the data completion model. The proposed MapT-STC algorithm incorporates double data sources to complete the incomplete matrix, and it considers the data correlation to achieve better results. In contrast, the previous DMF algorithm only focuses on the completion of a single incomplete matrix, which results in less accurate results when there are double data sources. In addition, MapT-STC uses a pretrained model as its initial model. Compared to the randomly initialized DMF, MapT-STC can train models with similar effects by fewer timeslots. In summary, the experimental

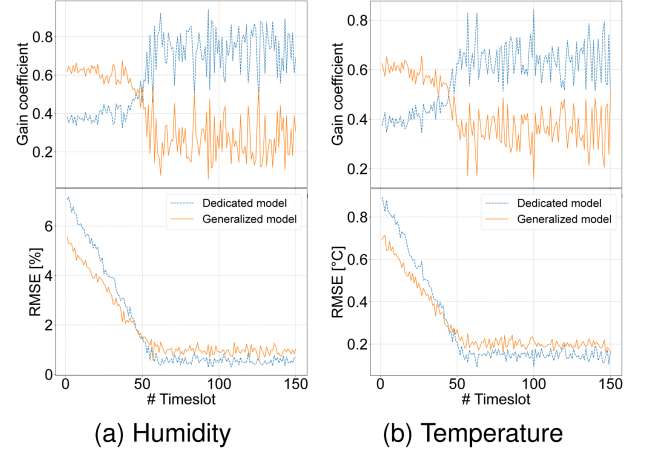


Fig. 15. The error and gain coefficient of the dedicated model and the generalized model under different timeslots over *Sensor-Scope*.

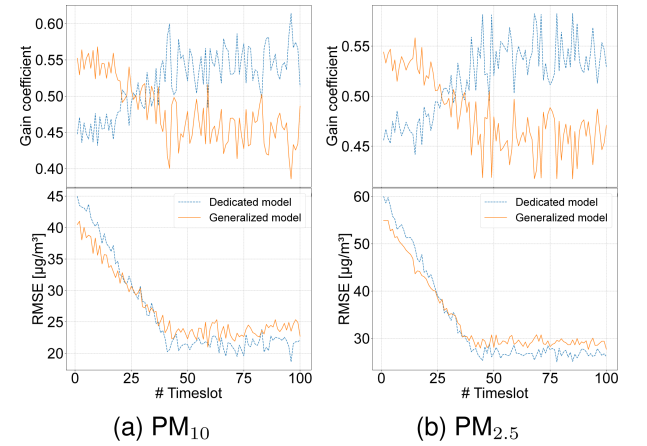
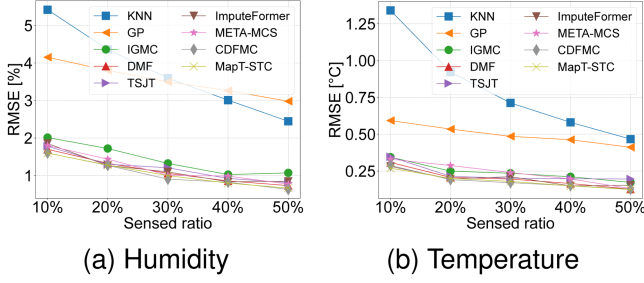
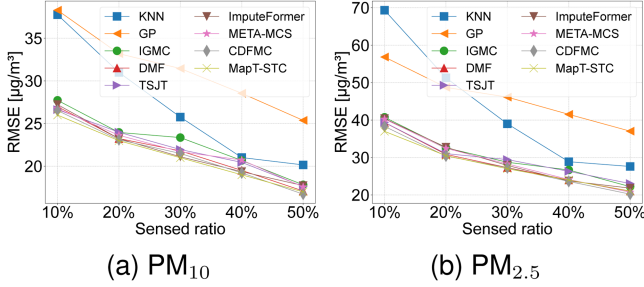


Fig. 16. The error and gain coefficient of the dedicated model and the generalized model under different timeslots over *U-Air*.

results demonstrate the superiority of the MapT-STC over the previous DMF and CDFMC in completing incomplete matrices with source datasets. The MapT-STC can achieve higher accuracy and require fewer numbers of timeslot.

E. Micro-Process Analysis (RQ3)

Our evaluation of both source dataset-based pretraining and target dataset-only self-training sub-modules across varying timeslot counts, as shown in Figs. 15 and 16, reveals that error (measured by *RMSE-Timeslot* curves) decreases for both components as timeslots increase. Initially, the source-dataset sub-module demonstrates lower error, but a clear turning point emerges where the target-dataset sub-module surpasses it, indicating that the target dataset's intrinsic completion capability becomes more effective than cross-dataset transfer during later training stages. Analysis of gain coefficient variations shows predominant contributions from the source-dataset sub-module pre-convergence, though dataset-specific patterns exist: for *Sensor-Scope*, coefficients increase then decrease, whereas for *U-Air* they remain initially stable before declining. These dynamic experimental results confirm MapT-STC's suitability for new crowdsensing tasks' initial phases, highlighting the


 Fig. 17. Inference accuracy under different sensed ratios over *Sensor-Scope*.

 Fig. 18. Inference accuracy under different sensed ratios over *U-Air*.

critical importance of assessing individual dataset contributions in training algorithm selection. Additionally, this experiment emphasizes that the completion effect of the target dataset itself may be better suited for some situations, providing better results than other datasets or a combination of multiple datasets. Further studies could explore other types of datasets and test the performance of more complex models to confirm these methods and findings.

F. Robustness Analysis (RQ4)

We conducted a robustness assessment of five methods across four tasks using two datasets, varying the sensed ratio from 10% to 50% to simulate differential sensed data availability. Results (Figs. 17 and 18) demonstrate that MapT-STC consistently achieves the lowest data completion error across all tasks and sensed ratios. While all methods exhibit decreasing error trends with higher sensed ratios, baseline methods (KNN, GP, IGMC, DMF, TSJT, ImputeFormer, META-MCS, CDFMC) persistently show higher errors. Comparative analysis reveals inherent limitations: KNN handles linear relationships but fails with complex data patterns; GP models non-linear relationships yet underperforms DNN-based approaches; IGMC's missing-data robustness depends on large-scale datasets, reducing applicability here; DMF, despite effectiveness on large datasets, is prone to overfitting and poor generalization. TSJT leverages transfer learning but struggles with domain shift under extreme sparsity (e.g., 10% sensed ratio), leading to unstable performance; ImputeFormer, though designed for low-rankness-induced spatiotemporal imputation and exhibiting strong signal-noise balance in principle, may over-smooth high-frequency details or amplify high-rank noise when handling irregularly sparse crowdsensing data; META-MCS adapts quickly to new tasks

TABLE VII
MINIMUM CONVERGENCE TIME FOR ABLATION VARIANTS OVER
SENSOR-SCOPE AND *U-AIR*

Method	Task			
	Humi.	Temp.	PM ₁₀	PM _{2.5}
Full MapT-STC	70 ± 4	62 ± 3	50 ± 3	50 ± 2
w/o CORAL	95 ± 6	89 ± 5	60 ± 4	64 ± 4
w/o Dedicated Model	82 ± 5	77 ± 4	54 ± 3	57 ± 3
w/o Kalman Fusion	80 ± 5	75 ± 4	52 ± 3	55 ± 3

via meta-learning but fails to capture global spatial-temporal correlations in large-scale scenarios, resulting in higher error fluctuations.

In contrast, MapT-STC integrates complementary strengths that collectively enable superior performance. First, it is able to model the correlation between data points, allowing it to better capture complex relationships in the data. Second, it is able to handle missing data effectively, enabling robust performance in scenarios with varying amounts of sensed data. Third, MapT-STC integrates correlative types of data, making it more suitable for dealing with practical sparse crowdsensing problems. Overall, the results of this experiment indicate that MapT-STC is a robust and effective method for matrix completion in various scenarios with varying amounts of sensed data. Its ability to model correlation and handle missing data, along with its computational efficiency, makes it a promising method for more and more real-world sparse crowdsensing applications.

G. Ablation Study (RQ5)

To verify the necessity of each component of MapT-STC, we conduct an ablation study focusing on the minimum convergence period (i.e., the number of timeslots required to achieve stable RMSE, defined as the point where error reduction plateaus within 5% of the final value). This study evaluates three variants: (1) MapT-STC without CORAL, where we replace CORAL-based feature alignment with direct transfer (i.e., applying the source-pretrained model directly to the target domain); (2) MapT-STC without the dedicated model, using only the generalized model; and (3) MapT-STC without Kalman-filter fusion, substituting it with simple averaging of the generalized and dedicated model outputs. Experiments are performed on the *U-Air* (PM_{2.5} and PM₁₀) and *Sensor-Scope* (temperature and humidity) datasets.

The framework overview in Fig. 4 illustrates the integration of CORAL, dual-path learning, and fusion modules. Removing any component disrupts this synergy, as validated below: We compare the full MapT-STC against the three variants. The convergence period is measured as the number of timeslots until the RMSE stabilizes, averaged over 5 runs. Results are summarized in Table VII.

Removing CORAL increases the convergence period by approximately 40% across tasks (e.g., from 62 to 89 for *Sensor-Scope* temperature). Without CORAL, the direct transfer approach fails to align feature distributions between source and

target domains, leading to higher initial errors and slower adaptation. This underscores CORAL's role in generating pseudo-historical data under cold start conditions with sparse sensed ratios, as described in Section IV. The alignment process is critical for reducing temporal dependencies and enabling effective domain adaptation. Omitting the dedicated model prolongs convergence by about 25% (e.g., *U-Air* PM_{2.5} from 50 to 57 timeslots). As shown in Fig. 7, the dedicated model captures target-specific patterns, while the generalized model handles cross-domain features. Without this dual-path balance, the generalized model alone struggles with target-domain nuances, increasing error volatility. This confirms the necessity of the dedicated model for rapid adaptation to local spatiotemporal dynamics. Replacing Kalman filtering with simple averaging extends convergence by 20% on average (e.g., *Sensor-Scope* humidity from 70 to 80 timeslots). Kalman fusion dynamically weights models based on uncertainty estimates (σ_G and σ_D), whereas averaging applies static weights, amplifying errors.

VII. CONCLUSION

This paper addresses the distinct and under-explored problem of one-shot data completion for cold-start tasks in sparse crowdsensing. Unlike few-shot transfer learning, our approach is designed for the regime where target-domain data is virtually absent. We introduce MapT-STC, whose novelty lies in a systematic integration that reshapes the pretraining-finetuning paradigm through extreme data sparsity, timeslot dependency, and uncertain spatiotemporal correlations. To overcome these limitations, we introduce MapT-STC—a novel framework integrating Deep Matrix Factorization (DMF) with Kalman Filtering (KF) to accurately estimate unsensed data. MapT-STC significantly extends our prior work [25] by enabling one-shot completion through mapping transfer learning, demonstrating broader applicability in data-scarce environments. MapT-STC fine-tunes a pre-trained model with minimal target data, enabling efficient adaptation to new sensing scenarios. By integrating DMF and KF, it effectively fuses source and target data for robust representation. Evaluations show it outperforms existing methods, achieving high-accuracy data completion with fewer timeslots and less training data. This makes it a promising solution for sparse crowdsensing, capable of handling incomplete data efficiently and accurately.

REFERENCES

- [1] L. Tan, H. Xiao, K. Yu, M. Aloqaily, and Y. Jararweh, "A blockchain-empowered crowdsourcing system for 5G-enabled smart cities," *Comput. Stand. Interfaces*, vol. 76, 2021, Art. no. 103517.
- [2] A. Nieto, A. Acien, and G. Fernandez, "Crowdsourcing analysis in 5G IoT: Cybersecurity threats and mitigation," *Mobile Netw. Appl.*, vol. 24, no. 3, pp. 881–889, 2019.
- [3] R. She, "Survey on incentive strategies for mobile crowdsensing system," in *Proc. IEEE 11th Int. Conf. Softw. Eng. Service Sci.*, 2020, pp. 511–514.
- [4] Y. Nakayama, "Network-side task allocation for mobile crowdsensing," in *Proc. IEEE 17th Annu. Consum. Commun. Netw. Conf.*, 2020, pp. 1–6.
- [5] J. Wu and E. Wang, *Mobile Crowdsourcing From Theory to Practice*. Berlin, Germany: Springer, 2023.
- [6] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: Current state and future challenges," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 32–39, Nov. 2011.
- [7] J. Bian, H. Xiong, Y. Fu, and S. K. Das, "CSWA: Aggregation-free spatial-temporal community sensing," in *Proc. AAAI Conf. Artif. Intell.*, 2018, pp. 2087–2094.
- [8] C. Xiang, P. Yang, C. Tian, H. Cai, and Y. Liu, "Calibrate without calibrating: An iterative approach in participatory sensing network," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 2, pp. 351–361, Feb. 2015.
- [9] C. Xiang et al., "CARM: Crowd-sensing accurate outdoor RSS maps with error-prone smartphone measurements," *IEEE Trans. Mobile Comput.*, vol. 15, no. 11, pp. 2669–2681, Nov. 2016.
- [10] X. Fan et al., "BuildSenSys: Reusing building sensing data for traffic prediction with cross-domain learning," *IEEE Trans. Mobile Comput.*, vol. 20, no. 6, pp. 2154–2171, Jun. 2021.
- [11] A. Ali, M. A. Qureshi, M. Shiraz, and A. Shamim, "Mobile crowd sensing based dynamic traffic efficiency framework for urban traffic congestion control," *Sustain. Comput. Informat. Syst.*, vol. 32, 2021, Art. no. 100608.
- [12] Y. Zhao, Z. Guan, and J. Zhang, "Return freight strategies and selling formats in e-commerce supply chain: The perspective of consumer fairness concerns and online shopping returns," *J. Retail. Consum. Serv.*, vol. 75, 2023, Art. no. 103493.
- [13] Z. Chen, Y. Lu, Z. Li, and X. Zhou, "Task grid-based urban environmental information release mechanism for mobile crowd sensing," *J. Electr. Comput. Eng.*, vol. 2022, pp. 1738660:1–1738660:10, 2022.
- [14] J. Liu, H. Cao, Q. Li, F. Cai, X. Du, and M. Guizani, "A large-scale concurrent data anonymous batch verification scheme for mobile healthcare crowd sensing," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1321–1330, Apr. 2019.
- [15] L. Wang et al., "CCS-TA: Quality-guaranteed online task allocation in compressive crowdsensing," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2015, pp. 683–694.
- [16] L. Wang et al., "SPACE-TA: Cost-effective task allocation exploiting intradata and interdata correlations in sparse crowdsensing," *ACM Trans. Intell. Syst. Technol.*, vol. 9, no. 2, pp. 20:1–20:28, 2018.
- [17] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in *Proc. Int. Conf. Learn. Representations*, Y. Bengio and Y. LeCun, Eds., 2014, pp. 1–14.
- [18] J. Fan and J. Cheng, "Matrix completion by deep matrix factorization," *Neural Netw.*, vol. 98, pp. 34–41, 2018.
- [19] E. Wang et al., "Deep learning-enabled sparse industrial crowdsensing and prediction," *IEEE Trans. Ind. Informat.*, vol. 17, no. 9, pp. 6170–6181, Sep. 2021.
- [20] M. Zhang and Y. Chen, "Inductive matrix completion based on graph neural networks," in *Proc. Int. Conf. Learn. Representations*, 2020, pp. 11738–11751.
- [21] K. Rezaee, S. J. Mousavirad, M. R. Khosravi, M. K. Moghimi, and M. Heidari, "An autonomous UAV-assisted distance-aware crowd sensing platform using deep shufflenet transfer learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 9404–9413, Jul. 2022.
- [22] J. Wang, C. Lan, C. Liu, Y. Ouyang, and T. Qin, "Generalizing to unseen domains: A survey on domain generalization," in *Proc. Int. Joint Conf. Artif. Intell.*, 2021, pp. 4627–4635.
- [23] A. Robey, G. J. Pappas, and H. Hassani, "Model-based domain generalization," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, M. Ranzato, A. Beygelzimer, Y. N. Dauphin, P. Liang, and J. W. Vaughan, Eds., 2021, pp. 20210–20229.
- [24] E. Wang, M. Zhang, Y. Yang, Y. Xu, and J. Wu, "Exploiting outlier value effects in sparse urban crowdsensing," in *Proc. IEEE/ACM 29th Int. Symp. Qual. Service*, 2021, pp. 1–10.
- [25] E. Wang, M. Zhang, B. Yang, Y. Xu, Z. Song, and Y. Yang, "Few-shot data completion for new tasks in sparse crowdsensing," in *Proc. IEEE Conf. Comput. Commun.*, 2024, pp. 1831–1840.
- [26] D. Zhang, L. Wang, H. Xiong, and B. Guo, "4W1H in mobile crowd sensing," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 42–48, Aug. 2014.
- [27] X. Lin, K. Wei, Z. Li, J. Chen, and T. Pei, "Aggregation-based dual heterogeneous task allocation in spatial crowdsourcing," *Front. Comput. Sci.*, vol. 18, no. 6, 2024, Art. no. 186605.
- [28] X. Zhu, Y. Luo, A. Liu, W. Tang, and M. Z. A. Bhuiyan, "A deep learning-based mobile crowdsensing scheme by predicting vehicle mobility," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4648–4659, Jul. 2021.
- [29] L. Wang, D. Zhang, Y. Wang, C. Chen, X. Han, and A. M'hamed, "Sparse mobile crowdsensing: Challenges and opportunities," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 161–167, Jul. 2016.
- [30] T. Liu, Y. Zhu, Y. Yang, and F. Ye, "Alc²: When active learning meets compressive crowdsensing for urban air pollution monitoring," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 9427–9438, Dec. 2019.

- [31] K. Xie, X. Li, X. Wang, G. Xie, J. Wen, and D. Zhang, "Active sparse mobile crowd sensing based on matrix completion," in *Proc. Int. Conf. Manag. Data*, 2019, pp. 195–210.
- [32] X. Chen et al., "Adaptive hybrid model-enabled sensing system (HMSS) for mobile fine-grained air pollution estimation," *IEEE Trans. Mob. Comput.*, vol. 21, no. 6, pp. 1927–1944, Jun. 2022.
- [33] L. Wang, D. Zhang, D. Yang, B. Y. Lim, X. Han, and X. Ma, "Sparse mobile crowdsensing with differential and distortion location privacy," *IEEE Trans. Inf. Forensics Secur.*, vol. 15, pp. 2735–2749, 2020.
- [34] J. Yoon, J. Jordon, and M. van der Schaar, "GAIN: Missing data imputation using generative adversarial nets," in *Proc. Int. Conf. Mach. Learn.*, J. G. Dy and A. Krause, Eds., 2018, pp. 5675–5684.
- [35] K. Xie et al., "Recover corrupted data in sensor networks: A matrix completion solution," *IEEE Trans. Mob. Comput.*, vol. 16, no. 5, pp. 1434–1448, May 2017.
- [36] W. Liu, Y. Yang, E. Wang, L. Wang, D. Zeglache, and D. Zhang, "Multi-dimensional urban sensing in sparse mobile crowdsensing," *IEEE Access*, vol. 7, pp. 82066–82079, 2019.
- [37] J. Han, H. Liu, H. Zhu, H. Xiong, and D. Dou, "Joint air quality and weather prediction based on multi-adversarial spatiotemporal networks," in *Proc. AAAI Conf. Artif. Intell.*, 2021, pp. 4081–4089.
- [38] E. Wang et al., "Outlier-concerned data completion exploiting intra- and inter-data correlations in sparse crowdsensing," *IEEE/ACM Trans. Netw.*, vol. 31, no. 2, pp. 648–663, Apr. 2023.
- [39] T. Nguyen and A. Grover, "Transformer neural processes: Uncertainty-aware meta learning via sequence modeling," in *Proc. Int. Conf. Mach. Learn.*, K. Chaudhuri, S. Jegelka, L. Song, C. Szepesvári, G. Niu, and S. Sabato, Eds., 2022, pp. 16569–16594.
- [40] Y. Du et al., "AdaRNN: Adaptive learning and forecasting of time series," in *Proc. 30th ACM Int. Conf. Inf. Knowl. Manage.*, G. Demartini, G. Zuccon, J. S. Culpepper, Z. Huang, and H. Tong, Eds., 2021, pp. 402–411.
- [41] A. Zadeh, P. P. Liang, N. Mazumder, S. Poria, E. Cambria, and L. Morency, "Memory fusion network for multi-view sequential learning," in *Proc. AAAI Conf. Artif. Intell.*, S. A. McIlraith and K. Q. Weinberger, Eds. AAAI Press, 2018, pp. 5634–5641.
- [42] Y. Zheng, F. Liu, and H. Hsieh, "U-air: When urban air quality inference meets Big Data," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2013, pp. 1436–1444.
- [43] F. Ingelrest, G. Barrenetxea, G. Schaefer, M. Vetterli, O. Couach, and M. Parlange, "SensorScope: Application-specific sensor network for environmental monitoring," *ACM Trans. Sens. Netw.*, vol. 6, no. 2, pp. 17:1–17:32, 2010.
- [44] F. Yin and F. Gunnarsson, "Distributed recursive Gaussian processes for RSS map applied to target tracking," *IEEE J. Sel. Top. Signal Process.*, vol. 11, no. 3, pp. 492–503, Apr. 2017.
- [45] H. Yen and C. Wang, "Cross-device Wi-Fi map fusion with Gaussian processes," *IEEE Trans. Mob. Comput.*, vol. 16, no. 1, pp. 44–57, Jan. 2017.
- [46] Y. Zhang, X. Wang, X. Yu, Z. Sun, K. Wang, and Y. Wang, "Drawing informative gradients from sources: A one-stage transfer learning framework for cross-city spatiotemporal forecasting," in *Proc. AAAI Conf. Artif. Intell.*, 2025, pp. 1147–1155.
- [47] T. Nie, G. Qin, W. Ma, Y. Mei, and J. Sun, "ImputeFormer: Low rankness-induced transformers for generalizable spatiotemporal imputation," in *Proc. 30th ACM SIGKDD Conf. Knowl. Discov. Data Mining*, 2024, pp. 2260–2271.
- [48] Z. Tian, E. Wang, W. Liu, B. Li, and F. Yang, "META-MCS: A meta-knowledge based multiple data inference framework," in *Proc. IEEE Conf. Comput. Commun.*, 2024, pp. 1351–1360.



sparse crowdsensing, spatiotemporal data mining, neural networks, and matrix completion.

Mijia Zhang (Member, IEEE) received the BE and ME degrees in electronic engineering from Harbin Engineering University, Harbin, China, in 2016 and 2019, respectively, and the PhD degree from in computer science from the College of Computer Science and Technology, Jilin University, Changchun, China, in 2023. He is currently a postdoctoral researcher with the College of Computer Science and Technology, Jilin University. He has also been selected for the "Dingxin Scholar" Support Program at Jilin University. His current work focuses on mobile computing,



Wei Li received the BS, MS, and PhD degrees from the field of Mathematics from the School of Mathematics, Jilin University, Changchun, China, in 1999, 2001, and 2009, respectively. She is currently a lecturer with the School of Business and Management, Jilin University. Her research interests include statistical machine learning, nonparametric estimation, stochastic processes, and their economic applications.



En Wang (Member, IEEE) received the BE degree in software engineering from Jilin University, Changchun, China, in 2011, and the ME and PhD degrees in computer science and technology from Jilin University, in 2013 and 2016, respectively. He was also a joint Ph.D. student with the Department of Computer and Information Science, Temple University, Philadelphia, PA, USA. He is currently a professor with the College of Computer Science and Technology, Jilin University. His current research interests include mobile computing, crowd intelligence, and data mining.



Jie Wu (Fellow, IEEE) is the director of the Center for Networked Computing and Laura H. Carnell professor with Temple University. He also serves as the director of International Affairs with the College of Science and Technology. He served as chair with the Department of Computer and Information Sciences from the summer of 2009 to the summer of 2016 and associate vice provost for International Affairs from the fall of 2015 to the summer of 2017. Prior to joining Temple University, he was a program Director at the National Science Foundation and was a distinguished professor at Florida Atlantic University. His current research interests include mobile computing and wireless networks, routing protocols, cloud and green computing, network trust and security, and social network applications. He regularly publishes in scholarly journals, conference proceedings, and books. He serves on several editorial boards, including *IEEE Transactions on Mobile Computing*, *IEEE Transactions on Service Computing*, *Journal of Parallel and Distributed Computing*, and *Journal of Computer Science and Technology*. He was general co-chair for IEEE MASS 2006, IEEE IPDPS 2008, IEEE ICDCS 2013, ACM MobiHoc 2014, ICPP 2016, and IEEE CNS 2016, as well as program co-chair for IEEE INFOCOM 2011 and CCF CNCC 2013. He was an IEEE Computer Society Distinguished Visitor, ACM Distinguished Speaker, and chair for the IEEE Technical Committee on Distributed Processing (TCDP). He is a CCF Distinguished Speaker. He is the recipient of the 2011 China Computer Federation (CCF) Overseas Outstanding Achievement Award.