

Change Detection in Partially Observed Large-Scale Traffic Network Data

Meng Zhao¹, Mostafa Reisi Gahrooei², and Mohammad Ilbeigi³

Abstract—Intelligent Transportation Systems generate an unprecedented amount of high-dimensional traffic data. The proper analysis of such data can transform traffic monitoring mechanisms. However, existing monitoring methods for detecting abrupt changes in traffic patterns have two limitations. First, they do not capture the spatiotemporal characteristics of traffic data and are not equipped with a built-in mechanism to handle missing observations. To address these limitations, this study proposes a dynamic, robust tensor completion method to monitor and detect changes in partially observed traffic data streams. The proposed method simultaneously completes and decomposes the partially observed data into a sum of a low-rank tensor that captures the spatiotemporal patterns and a sparse tensor that captures anomalies. Subsequently, the proposed method defines a statistic monitored by an exponentially weighted moving average control chart to detect abrupt temporal changes. The performance of the proposed method is evaluated by simulation and case studies. The simulation results indicate the proposed method outperforms all benchmarks. It can also detect changes more than twice as fast as other benchmarks in terms of average run length in most scenarios. The proposed method is also applied to the traffic data in New York City to evaluate its performance in detecting unusual traffic patterns when Hurricane Sandy hit the city. The experimental results demonstrated the superiority of the proposed method in quickly detecting unusual changes at both network and road segment levels. Particularly, the proposed method detects changes in traffic patterns approximately twelve hours earlier than the next best alternative benchmark method.

Index Terms—High-dimensional incomplete data streams, robust tensor completion, statistical monitoring.

I. INTRODUCTION

WITH the rapid advancement of sensing technologies in Intelligent Transportation Systems (ITS), an unprecedentedly large amount of spatiotemporal traffic data has become available. Major cities have tens of thousands of roads, and various ITS technologies, such as vehicles equipped with Global Positioning Systems (GPS), loop detectors, and ultrasound Doppler radars, continuously generate traffic data from many of these roads. Such big data streams offer opportunities for more effective, precise, and efficient traffic monitoring

mechanisms. Comprehensive traffic monitoring systems that cover large-scale road networks can help decision-makers detect abnormal traffic patterns quickly, plan to mitigate congestion proactively, and support emergency operations in the aftermath of extreme events.

Efficiently analyzing high-dimensional (HD) traffic data requires transformative changes in state-of-the-art methods for traffic monitoring. HD traffic data includes a large number of potentially highly correlated features, representing the current state of the traffic network. More specifically, at each time, the state of the traffic network is explained by thousands of observations, each indicating the traffic condition of a road segment, that may correlate with each other. Each observation can be considered as a feature of the traffic network at a given time. These features should be viewed collectively as modeling and monitoring them individually often results in loss of knowledge and a significant number of false alarms. As it is described in more detail in this section, the existing solutions for traffic monitoring are not scalable to handle HD data and may not effectively capture their complex characteristics. More specifically, designing such methods is complicated due to two challenges. The first challenge is to capture the complex characteristics of HD traffic data, including high dimensionality, spatiotemporal correlation structure, and non-stationarity [1]. A monitoring method with these capabilities can quickly detect abrupt changes with fewer false alarms. The second challenge is to handle incomplete and partially observed traffic data. Traffic datasets, even when captured by advanced technologies, are often incomplete. Permanent traffic counters are not available on all roads, and vehicles equipped with GPS devices may not travel on all roads at all times. As a result, traffic data is only partially observable at each point in time. Occasional sensor failures and technical issues in data transmissions may also result in further missing observations [2], [3]. Therefore, effective traffic monitoring methods require the ability to detect unusual patterns in HD spatiotemporal traffic data streams that contain considerable missing values.

A. Existing Methods for Traffic Monitoring and Their Limitations

Overall, previous studies in traffic monitoring can be categorized into three main groups. The first group of studies uses non-anomalous historical data to create short-term forecasting models with confidence intervals and alarm for an abnormal pattern when a prediction falls outside the intervals [4], [5], [6], [7]. The second group of studies develops classification

Manuscript received 23 February 2024; revised 17 June 2024; accepted 4 August 2024. This work was supported in part by the US National Science Foundation under Grant CMMI-2027025 and Grant CMMI-2026795. The Associate Editor for this article was F. Xia. (Corresponding author: Mostafa Reisi Gahrooei.)

Meng Zhao and Mostafa Reisi Gahrooei are with the Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: m.zhao@ufl.edu; mreisigahrooei@ufl.edu).

Mohammad Ilbeigi is with the Department of Civil, Environmental, and Ocean Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: milbeigi@stevens.edu).

Digital Object Identifier 10.1109/TITS.2024.3440836

1558-0016 © 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See <https://www.ieee.org/publications/rights/index.html> for more information.

models to recognize unusual patterns using various machine learning methods such as artificial neural networks [8], [9], [10], support vector machines [11], [12], [13], and random forests [14], [15]. Both predictive and classification methods quickly lose their statistical power when applied to networks with a large number of roads and lead to a significant number of false alarms. Therefore, they can effectively monitor only a very limited number of roads and are not scalable to handle HD data.

The third group of studies proposes network-wide shift detection methods. Typically, these methods summarize HD traffic data from a road network into a monitoring statistic as a proxy to explain the overall network-level patterns. The monitoring statistics can be defined using different approaches, including based on a probability distribution that explains variations of traffic speed on the roads [16] or dynamic network topological characteristics (e.g., network betweenness centrality index) [17]. Although these methods can monitor a large network as a whole, they cannot capture spatiotemporal correlations and interdependencies among roads. This limitation prevents them from explaining what part of a network is experiencing abnormal traffic when a change is detected at the network level.

In addition, none of the existing solutions offer a systematic approach to handle partially observed data during the monitoring process. The fundamental limitations in the existing traffic monitoring methods to handle HD spatiotemporal traffic data call for the development of new monitoring methods.

Although the existing methods for traffic monitoring are subject to essential limitations in handling partially observed HD data, recent advancements in other domains, such as manufacturing, focused on developing monitoring mechanisms that are able to capture complex characteristics in incomplete and partially observed HD data may seem viable potential solutions to be adopted in the context of traffic. Therefore, in the following two subsections, we review state-of-the-art methods for HD data monitoring developed in other areas and discuss their critical limitations for potential applications in traffic monitoring.

B. Existing Methods for HD Data Monitoring in Other Domains

A group of studies developed for monitoring HD data use LASSO estimators to create Lasso-based control charts and detect unusual shifts in the data streams by declaring the non-zero values as anomalies [18], [19], [20]. While these methods are able to handle HD data, they cannot address the complex spatiotemporal characteristics of HD traffic datasets, which may result in large number of false alarms.

Some other studies use dimensionality reduction techniques to embed the HD data onto a lower dimensional space where anomalies can be identified [21], [22]. For instance, Liu [23] used T^2 and Q charts for reducing dimensions and monitoring spatiotemporal data. Paynabar et al. [24] proposed a functional principle complement-based control chart to monitor multivariate functional data. Tensor (multi-dimensional arrays) analysis has also been used for HD data

dimension reduction and monitoring. Yan et al. [1] provided a comprehensive introduction to the use of low-rank tensor decomposition (LRTD) and dimension reduction methods, including multilinear principal component analysis (MPCA) [25], uncorrelated multilinear principal component analysis (UMPCA) [26], and tensor rank-one decomposition (TROD) for process monitoring. Additionally, Qian et al. [27] proposed a CP decomposition to extract discriminant features from signal streams, and Hou et al. [28] designed a three-order Tucker decomposition and reconstruction-based change detection framework.

The methods designed based on dimension reduction are subject to two main drawbacks for implementation in a traffic context. First, they may not fully capture the spatiotemporal characteristics of the data, particularly when techniques such as unfolded PCA are used. This limitation is because they may destroy the inherent correlations and interactions within the data by unfolding the HD data (e.g., an image) into a vector. Second, the extracted features may not capture a change as they are extracted by combining the global patterns that can mask smaller abnormalities. Another group of studies proposed methods that decompose the data into a smooth (or low-rank) background and sparse outliers. For example, Candès et al. [29] proposed a robust PCA model (RPCA) to decompose the observed data into a low-rank background component and a sparse component. This model has drawn a lot of attention in anomaly detection in numerous applications [30], [31], [32]. While these methods can exploit the spatial characteristics of the data, they do not account for the temporal patterns in the data stream.

C. Existing Methods for Handling Partially Observed Data in Monitoring

A typical solution to manage partially observed data is to develop data imputation methods. These methods first complete the dataset and then apply the monitoring processes. Classic imputation methods for traffic data use three main approaches, including prediction [33], [34], [35], [36], [37], [38], [39], [40], interpolation [41], [42], [43], and statistical learning [44], [45]. These methods have inherent limitations in completing HD spatiotemporal data [2], [46].

More advanced data imputation approaches apply decomposition-based techniques, such as low-rank tensor completion methods [47], [48]. These methods reconstruct data profiles from a low-dimensional representation of the collected data using nuclear norm approximation [49] or by imposing a predefined decomposition form [50], [51]. Despite the superiority of these methods against classic data imputation methods in HD data, they may not fully capture the clustered structure of the HD spatiotemporal data. To address this limitation, a recent study [2] extended PARATUCK2 matrix decomposition to develop a novel data imputation method for HD traffic data with spatial and temporal clusters. Despite the considerable advancements in data imputation methods, employing these solutions leads to a two-step process for traffic monitoring in which the data streams must first be completed and then monitored.

Furthermore, the performance of the imputation methods and the accuracy of the estimated values for missing observations may critically affect the reliability of the monitoring process. A data monitoring process that can inherently manage missing values and partially observed large-scale traffic data can improve the efficiency of traffic monitoring mechanisms.

D. Objective and Overview of the Proposed Method

Motivated by the limitations in the existing literature, the objective of this article is to develop a new method for monitoring and change detection in partially observed HD traffic network data streams. This study contributes to the body of knowledge by proposing a new monitoring method with two novel capabilities:

- The proposed method is equipped with a built-in data imputation approach that can handle simultaneously partially observed traffic data as an integrated part of the monitoring process.
- The proposed method can detect unusual traffic patterns at both network and road levels of granularity.

The proposed method first estimates the missing values of the HD data streams by using both spatial and temporal information. This is achieved by constructing a new objective function that utilizes the dynamic information from past data. Second, it separates the sparse components (outliers) from the natural observation patterns and monitors the sparse components for change detection. More specifically, we develop a spatiotemporal robust tensor completion approach combined with an exponentially weighted moving average (EWMA) control chart. To capture the temporal patterns of data streams and enhance the precision of missing data completion and anomaly detection, two Frobenius norm-based penalties are introduced to minimize the difference between two adjacent observations (each represented by a tensor) in the data stream. To detect a change in the data stream, the L_1 norm of anomaly tensors is used as the monitoring statistic, monitored by an EWMA control chart. For convenience, we name the proposed approach, robust spatiotemporal tensor completion-based monitoring (RSTCM), in the following sections.

The remainder of this article is organized as follows: In Section II, we introduce notations related to our defined problem and the background of the RTC problem. Section III presents the formulation of the proposed RSTCM model and the algorithmic approach to model estimation. In Section IV, a simulation analysis is conducted to evaluate the performance of the proposed method. Next, the proposed method is applied to New York City (NYC) historical traffic data to examine its performance in detecting abnormal traffic patterns due to Hurricane Sandy in 2012 in Section V. Finally, Section VI summarizes the proposed method and concludes the primary contributions of this article to the body of knowledge.

II. MULTILINEAR ALGEBRA AND ROBUST TENSOR COMPLETION

In this section, tensor notations and basic definitions used throughout this article are introduced. In addition, the robust tensor recovery and completion problems are reviewed.

A. Tensor Basics and Multilinear Algebra

Throughout this paper, we denote scalars, vectors, matrices, and tensors by lowercase or capital letters (a or A), boldface lowercase letters ($\mathbf{a}$), boldface capital letters ($\mathbf{A}$), and Euler script letters ($\mathcal{A}$), respectively. A fiber of the tensor $\mathcal{A}$ is a column vector defined by fixing all but one index. The unfolding of the tensor $\mathcal{A}$ in the i -th mode results in a matrix $\mathbf{A}_{(i)}$ by augmenting the i -th mode fibers as its columns. This process is also called matricization. The inner product of two matrices $\mathbf{A}$ and $\mathbf{B}$ is defined as $\langle \mathbf{A}, \mathbf{B} \rangle = \text{Tr}(\mathbf{A}^T \mathbf{B})$ and the inner product of two tensors $\mathcal{A}$ and $\mathcal{B}$ is defined as $\langle \mathcal{A}, \mathcal{B} \rangle = \langle \mathbf{A}_{(j)}, \mathbf{B}_{(j)} \rangle$. The symbols $\otimes$ and $\odot$ denote the Kronecker and Khatri-Rao products, respectively.

For a N -th order tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, its Frobenius norm is defined as the square root of the sum of the squares of its elements, i.e., $\|\mathcal{A}\|_F = \|\mathbf{A}_{(i)}\|_F = \sqrt{\mathbf{A}_{(i)}^T \mathbf{A}_{(i)}}$. Its L_1 norm is defined as the sum of the absolute value of its entries, denoted by $\|\mathcal{A}\|_1 = \sum_{i_1 i_2 \dots i_N} |\mathcal{A}_{i_1 i_2 \dots i_N}|$. The nuclear norm of $\mathcal{A}$ is defined as the sum of nuclear norms of the tensor $\mathcal{A}$ unfoldings in all N modes, denoted by $\|\mathcal{A}\|_* = \sum_i \|\mathbf{A}_{(i)}\|_*$, where $\|\mathbf{A}_{(i)}\|_* = \sum_j \sigma_j$ and σ_j is the j^{th} singular value of the matrix $\mathbf{A}_{(i)}$ via the singular value decomposition (SVD). The CANDECOMP/PARAFAC (CP) decomposition factorizes the tensor $\mathcal{A}$ into a sum of rank-one tensors, which is denoted as $\mathcal{A} \approx [\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \dots, \mathbf{A}^{(N)}] \equiv \sum_{r=1}^R \mathbf{a}_r^{(1)} \circ \mathbf{a}_r^{(2)} \circ \dots \circ \mathbf{a}_r^{(N)}$, where the notation $\circ$ is the outer product operation.

B. Review of Robust Tensor Recovery and Completion

This subsection summarizes the robust tensor recovery and completion formulations, which provides a basis for our proposed approach. The robust tensor recovery (RTR) attempts to recover a low-rank tensor from a tensor that is corrupted by sparse outliers. That is, RTR will decompose a higher-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ into a summation of a low-rank tensor $\mathcal{L} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ and a sparse tensor $\mathcal{S} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, so that $\mathcal{L} + \mathcal{S}$ matches $\mathcal{X}$. The locations and the values of outliers (i.e., non-zero elements of the sparse tensor $\mathcal{S}$) are not known apriori. To achieve this decomposition, $\text{rank}(\mathcal{L}) + \lambda \|\mathcal{S}\|_1$ is minimized under the constraint that $\mathcal{X} = \mathcal{L} + \mathcal{S}$. However, this problem is intractable because identifying the rank of a tensor is an NP-hard problem. Therefore, various convex relaxations of this problem have been proposed that use the nuclear norm of a tensor as a convex proxy to the rank penalty [52], [53], [54]. The hyper-parameter λ is often selected empirically. Particularly, [53] provided evidence that the tuning parameter λ can be set to $1/\sqrt{\max(I_1, I_2)I_3}$ for a third-order tensor with dimensions I_1 , I_2 , and I_3 to guarantee the exact recovery.

A tensor completion (TC) problem uses the low-rank properties of a tensor to impute missing values of the tensor. To be more precise, let $\Omega \subset \{1, \dots, I_1\} \times \{1, \dots, I_2\} \times \dots \times \{1, \dots, I_N\}$ denotes the set of observed entries of $\mathcal{X}$ and $\mathcal{X}^\Omega$ represent the projection of tensor $\mathcal{X}$ onto the subspace supported on Ω , defined as follows

$$\mathcal{X}^\Omega = \begin{cases} \mathcal{X}_{i_1 i_2 \dots i_N}, & (i_1, i_2, \dots, i_N) \in \Omega, \\ 0, & \text{otherwise.} \end{cases}$$

The TC problem then estimates the complete tensor $\mathcal{Z}$ by minimizing $\text{rank}(\mathcal{Z})$ such that $\mathcal{Z}^\Omega = \mathcal{X}^\Omega$. Similar to the TR problem, the rank penalty is often replaced by a convex proxy such as nuclear norm [55].

The robust tensor completion (RTC) combines tensor recovery and completion to recover a low-rank tensor from a corrupted and partially observed tensor. The objective of the RTC model is to determine a low-rank tensor $\mathcal{L}$ and a sparse tensor $\mathcal{S}$ so that $\mathcal{Z} = \mathcal{L} + \mathcal{S}$ represents the imputed version of partially observed tensor $\mathcal{X}$. Tensors $\mathcal{L}$ and $\mathcal{S}$ are estimated by minimizing $\text{rank}(\mathcal{L}) + \lambda \|\mathcal{S}\|_1$, subject to $\mathcal{X}^\Omega = (\mathcal{L} + \mathcal{S})^\Omega$. The rank penalty has been approximated by various other norms to achieve a computationally tractable problem [56], [57], [58], [59]. Although existing RTC techniques can complete partial observations and identify sparse components effectively [3], they are not designed to capture the temporal patterns of tensor data streams. Therefore, they are not directly applicable to the effective monitoring and change detection of partially observed HD data streams.

III. PROPOSED RSTCM MODEL FOR TRAFFIC MONITORING AND CHANGE DETECTION

In this section, we present the proposed RSTCM model based on the RTC formulation described in Section II-B to capture the temporal characteristics of incomplete HD traffic data streams for monitoring and change detection.

A. Problem Formulation

We consider and monitor HD traffic data as a stream of tensors $\{\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_t, \dots\}$ for detecting abrupt and unexpected changes. Here, $\mathcal{X}_t \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is a partially observed tensor, representing traffic data acquired at a time window t . Each tensor may include outliers that must be detected in order to efficiently and effectively detect statistically significant changes in traffic patterns. One potential approach to monitoring a stream of partially observed tensors is directly applying RTC to each tensor at time t to complete tensors and extracting outliers to be monitored for change detection. However, this approach ignores the temporal correlation structure of the traffic data stream when performing tensor completion and outlier detection. In other words, This issue may lead to excessive false alarms as natural patterns of data may be detected as abrupt changes. Our goal is to incorporate the temporal relationship among tensors when performing tensor completion and outlier detection in our proposed method. This allows for separating the natural temporal patterns from the abrupt changes and consequently reducing the false alarms.

Let $\mathcal{Z}_t$ denote the completed version of tensor $\mathcal{X}_t$. At each time t , we decompose $\mathcal{Z}_t$ as a summation of a low-rank tensor $\mathcal{L}_t$ and a sparse tensor of outliers $\mathcal{S}_t$ while considering the temporal correlation between current tensor data and previously observed ones. In order to capture the temporal correlation, we include a penalty term that penalizes deviation of $\mathcal{L}_t$ from $\mathcal{L}_{t-1}$ enforcing them to remain within each others proximity. Therefore, we solve the following problem to estimate $\mathcal{L}_t$ and

$\mathcal{S}_t$:

$$\begin{aligned} \min_{\mathcal{L}_t, \mathcal{S}_t} \quad & \|\mathcal{L}_t\|_* + \lambda \|\mathcal{S}_t\|_1 + \frac{\alpha}{2} \|\mathcal{L}_t - \mathcal{L}_{t-1}\|_F^2, \\ \text{s.t.} \quad & \mathcal{Z}_t = \mathcal{L}_t + \mathcal{S}_t. \end{aligned} \quad (1)$$

The challenge in solving (1) is that $\mathcal{Z}_t$ is not known; instead, we only observe $\mathcal{X}_t$ that contains missing entries outside the set Ω_t , where $\Omega_t \subset \{1, \dots, I_1\} \times \{1, \dots, I_2\} \times \dots \times \{1, \dots, I_N\}$. To address this challenge following [3], we first introduce a compensation tensor $\mathcal{M}_t \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ whose entries are zero in the support set Ω_t and takes non-zero values outside Ω_t . Let $\bar{\Omega}_t \subset \{1, \dots, I_1\} \times \{1, \dots, I_2\} \times \dots \times \{1, \dots, I_N\}$ denote the complement set of Ω_t to denote the index set where elements in $\mathcal{X}_t$ are not observed. Then, we define

$$\mathcal{M}_t = \begin{cases} \mathcal{M}_{i_1 i_2 \dots i_N}, & (i_1, i_2, \dots, i_N) \in \bar{\Omega}_t, \\ 0 & \text{otherwise.} \end{cases}$$

Using this compensation tensor, we update the problem formulation to estimate $\mathcal{L}_t$, $\mathcal{S}_t$, and $\mathcal{M}_t$ so that $\mathcal{X}_t^{\Omega_t} = \mathcal{L}_t + \mathcal{S}_t + \mathcal{M}_t$. In this formulation, the compensation tensor $\mathcal{M}_t$ cancels $\mathcal{L}_t + \mathcal{S}_t$ to produce the zero values in $\mathcal{X}_t^{\Omega_t}$. That is, $\mathcal{M}_t$ contains the negative of estimated missing values in $\mathcal{X}_t$. When estimating these missing values, not only do we use the spatial characteristics of $\mathcal{X}_t$, but we also consider the temporal relation between the current and previously observed tensors. For this purpose, we encourage $\mathcal{M}_t$ to remain within the proximity of $\mathcal{L}_{t-1}^{\bar{\Omega}_t}$.

Thus, we formulate our proposed objective function as

$$\begin{aligned} \min_{\mathcal{L}_t, \mathcal{S}_t, \mathcal{M}_t} \quad & \|\mathcal{L}_t\|_* + \lambda \|\mathcal{S}_t\|_1 + \frac{\alpha}{2} \|\mathcal{L}_t - \mathcal{L}_{t-1}\|_F^2 \\ & + \frac{\beta}{2} \|\mathcal{L}_{t-1}^{\bar{\Omega}_t} + \mathcal{M}_t\|_F^2, \\ \text{s.t.} \quad & \mathcal{X}_t^{\Omega_t} = \mathcal{L}_t + \mathcal{S}_t + \mathcal{M}_t, \end{aligned} \quad (2)$$

where the first term imposes a low-rank penalty to recover the underlying low-rank tensor, and the second term imposes sparsity on the outlier tensor. The third term includes a penalty so that the entries of low-rank tensors at time t and $t-1$ to be similar due to the temporal characteristics of the data streams, and similarly, the fourth term encourages the entries in $\mathcal{M}_t$ at time t to be similar to the corresponding entries of the projected low-rank tensor $\mathcal{L}_{t-1}$ on $\bar{\Omega}_t$. Please note that the low-rank tensor $\mathcal{L}_{t-1}$ is assumed to be known at time t , and $\lambda \geq 0$, $\alpha \geq 0$, and $\beta \geq 0$ are three hyperparameters that need to be tuned to balance the weighted losses in the objective function.

To solve problem (2), we use the alternating direction method of multipliers (ADMM), which can solve constrained convex optimization problems with separable objective functions (into differentiable and non-differentiable terms) efficiently. While ADMM provides a framework to approach this problem, it does not directly solve the problem. Instead, it translates the problem into other optimization problems, as we discuss later on. When applying ADMM, we first write the corresponding augmented Lagrangian function for

problem (2), where we simplify $L_\rho(\mathcal{L}_t, \mathcal{S}_t, \mathcal{M}_t, \mathcal{Y}_t)$ to L_ρ

$$\begin{aligned} L_\rho = & \|\mathcal{L}_t\|_* + \lambda \|\mathcal{S}_t\|_1 + \frac{\alpha}{2} \|\mathcal{L}_t - \mathcal{L}_{t-1}\|_F^2 \\ & + \frac{\beta}{2} \|\mathcal{L}_{t-1}^{\tilde{\Omega}_t} + \mathcal{M}_t\|_F^2 \\ & + \langle \mathcal{X}_t - \mathcal{L}_t - \mathcal{S}_t - \mathcal{M}_t, \mathcal{Y}_t \rangle \\ & + \frac{\rho}{2} \|\mathcal{X}_t - \mathcal{L}_t - \mathcal{S}_t - \mathcal{M}_t\|_F^2, \end{aligned} \quad (3)$$

where $\mathcal{Y}_t$ are the Lagrange multipliers or dual variables and ρ is a positive scalar. Using the property of the Frobenius norm, $\|\mathcal{A} + \mathcal{B}\|_F^2 = \|\mathcal{A}\|_F^2 + \|\mathcal{B}\|_F^2 + 2\langle \mathcal{A}, \mathcal{B} \rangle$, minimizing (3) is equivalent to minimizing the following objective function,

$$\begin{aligned} L_\rho = & \|\mathcal{L}_t\|_* + \lambda \|\mathcal{S}_t\|_1 + \frac{\beta}{2} \|\mathcal{L}_{t-1}^{\tilde{\Omega}_t} + \mathcal{M}_t\|_F^2 \\ & + \frac{\gamma}{2} \|\mathcal{L}_t - \frac{1}{\gamma}(\alpha \mathcal{L}_{t-1} + \rho(\mathcal{X}_t - \mathcal{S}_t - \mathcal{M}_t + \mathcal{Y}_t/\rho))\|_F^2, \end{aligned} \quad (4)$$

where $\gamma = \alpha + \rho$. Minimizing (4) cannot be directly achieved due to the definition of the tensor nuclear norm. Let us denote $\mathcal{L}_{t-1}^{\tilde{\Omega}_t}$ by $\mathcal{H}$ and drop the time index t for simplicity of notation. Also, set $\mathcal{E} = \mathcal{X} - \mathcal{S} - \mathcal{M} + \mathcal{Y}/\rho$. Then, the matricized version of the above objective function is as follows:

$$\begin{aligned} L_\rho = & \frac{1}{N} \sum_{i=1}^N \left(\|\mathbf{L}_{(i)}\|_* + \lambda \|\mathbf{S}_{(i)}\|_1 + \frac{\beta}{2} \|\mathbf{H}_{(i)} + \mathbf{M}_{(i)}\|_F^2 \right. \\ & \left. + \frac{\gamma}{2} \|\mathbf{L}_{(i)} - \frac{1}{\gamma}(\alpha \mathbf{H}_{(i)} + \rho \mathbf{E}_{(i)})\|_F^2 \right), \end{aligned} \quad (5)$$

where $\mathbf{E}_{(i)}$ is the folded (matricized) version of $\mathcal{E}$ along the mode i . Because the folded tensors, $\mathbf{L}_{(i)}$, along different modes, are coupled and share the same elements, we cannot minimize this objective function separately for each folded tensor. To address this issue, we define N auxiliary tensors, $\mathcal{R}_i = \mathcal{L}$ ($i = 1, \dots, N$). We use these auxiliary tensors to decouple the problem by replacing the folded tensor $\mathbf{L}_{(i)}$ with $\mathbf{R}_{i(i)}$ as follows to obtain:

$$\begin{aligned} L_\rho = & \frac{1}{N} \sum_{i=1}^N \left(\|\mathbf{R}_{i(i)}\|_* + \lambda \|\mathbf{S}_{(i)}\|_1 + \frac{\beta}{2} \|\mathbf{H}_{(i)} + \mathbf{M}_{(i)}\|_F^2 \right. \\ & \left. + \frac{\gamma}{2} \|\mathbf{R}_{i(i)} - \frac{1}{\gamma}(\alpha \mathbf{H}_{(i)} + \rho \mathbf{E}_{(i)})\|_F^2 \right), \end{aligned} \quad (6)$$

subject to $\mathcal{R}_i = \mathcal{L}$ for $i = 1, \dots, N$. Adding these constraints into the above objective function gives a new augmented lagrangian function as follows:

$$\begin{aligned} L_\rho = & \frac{1}{N} \sum_{i=1}^N \left(\|\mathbf{R}_{i(i)}\|_* + \lambda \|\mathbf{S}_{(i)}\|_1 \right. \\ & + \frac{\beta}{2} \|\mathbf{H}_{(i)} + \mathbf{M}_{(i)}\|_F^2 \\ & + \frac{\gamma}{2} \|\mathbf{R}_{i(i)} - \frac{1}{\gamma}(\alpha \mathbf{H}_{(i)} + \rho \mathbf{E}_{(i)})\|_F^2 \\ & \left. + \langle \mathcal{C}_i, \mathcal{R}_i - \mathcal{L} \rangle + \frac{\zeta}{2} \|\mathcal{R}_i - \mathcal{L}\|_F^2 \right), \end{aligned} \quad (7)$$

where $\mathcal{C}_i$ ($i = 1, \dots, N$) is the tensor of dual variables and ζ is a hyperparameter. Then, the ADMM algorithm iteratively

updates the set of variables $\mathcal{R}_i, \mathcal{S}, \mathcal{M}, \mathcal{Y}, \mathcal{C}_i$, and $\mathcal{L}$. More specifically, when estimating one of the variables, we fix others similar to the block coordinate algorithms [60], [61]. Accordingly, updating $\mathcal{R}_i$ ($i = 1, \dots, N$) is to solve the following problem

$$\begin{aligned} \min_{\mathbf{R}_{i(i)}} & \|\mathbf{R}_{i(i)}\|_* + \frac{\gamma}{2} \|\mathbf{R}_{i(i)} - \frac{1}{\gamma}(\alpha \mathbf{H}_{(i)} + \rho \mathbf{E}_{(i)})\|_F^2 \\ & + \langle \mathcal{C}_{i(i)}, \mathbf{R}_{i(i)} - \mathbf{L}_{(i)} \rangle + \frac{\zeta}{2} \|\mathbf{R}_{i(i)} - \mathbf{L}_{(i)}\|_F^2, \end{aligned} \quad (8)$$

that can be re-written as

$$\begin{aligned} \min_{\mathbf{R}_{i(i)}} & \|\mathbf{R}_{i(i)}\|_* \\ & + \frac{\omega}{2} \left\| \mathbf{R}_{i(i)} - \frac{1}{\omega} \left(\mathbf{H}_{(i)} + \rho \mathbf{E}_{(i)} + \zeta \left(\mathbf{L}_{(i)} - \frac{\mathbf{C}_{i(i)}}{\zeta} \right) \right) \right\|_F^2, \end{aligned} \quad (9)$$

where $\omega = \gamma + \zeta$. Next, updating the global variable $\mathcal{L}$ is to solve the following problem:

$$\min_{\mathcal{L}} \sum_{i=1}^N \langle \mathcal{C}_i, \mathcal{R}_i - \mathcal{L} \rangle + \frac{\zeta}{2} \|\mathcal{R}_i - \mathcal{L}\|_F^2 \quad (10)$$

Updating $\mathcal{S}_t$ is to solve the following problem

$$\min_{\mathcal{S}} \lambda \|\mathcal{S}\|_1 + \frac{\gamma}{2} \|\mathcal{L} - \frac{1}{\gamma}(\alpha \mathcal{H} + \rho \mathcal{E})\|_F^2. \quad (11)$$

Updating $\mathcal{M}$ is to solve the following problem

$$\min_{\mathcal{M}} \frac{\beta}{2} \|\mathcal{Q}\|_F^2 + \frac{\gamma}{2} \|\mathcal{L} - \frac{1}{\gamma}(\alpha \mathcal{H} + \rho \mathcal{E})\|_F^2, \quad (12)$$

where $\mathcal{Q} = \mathcal{H} + \mathcal{M}$. Finally, updating the Lagrange multipliers $\mathcal{Y}$ and $\mathcal{C}_i$ ($i = 1, \dots, N$) are as follows

$$\mathcal{Y}^{k+1} = \mathcal{Y}^k + \rho(\mathcal{X} - \mathcal{L} - \mathcal{S} - \mathcal{M}), \quad (13)$$

$$\mathcal{C}_i^{k+1} = \mathcal{C}_i^k + \zeta(\mathcal{R}_i - \mathcal{L}), \quad (14)$$

where k denotes the iteration index. Before we provide solutions to the above problems, let us define the soft thresholding operation $\Theta_u(\mathcal{X})$ for the tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ as

$$\Theta_u(\mathcal{X})_{i_1 i_2 \dots i_N} = (\mathcal{X}_{i_1 i_2 \dots i_N} - u)_+ - (-\mathcal{X}_{i_1 i_2 \dots i_N} - u)_+.$$

Also, for the matrix $\mathbf{X} \in \mathbb{R}^{M \times N}$, with the singular value decomposition $\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, the hard thresholding operation $\Phi_u(\mathbf{X})$ is defined as

$$\Phi_u(\mathbf{X}) = \mathbf{U}\Theta_u(\mathbf{S})\mathbf{V}^T.$$

The following propositions provide the solution to these above optimization problems to update each variable.

Proposition 1: Given $\mathcal{S}^k, \mathcal{M}^k, \mathcal{Y}^k, \mathcal{L}^k$, and $\mathcal{C}_i^k$, problem (9) has a closed form solution as follows:

$$\mathbf{R}_{i(i)}^{k+1} = \Phi_{\frac{1}{\omega}} \left(\frac{1}{\omega} \left(\mathbf{H}_{(i)} + \rho \mathbf{E}_{(i)} + \zeta \left(\mathbf{L}_{(i)} - \frac{\mathbf{C}_{i(i)}}{\zeta} \right) \right) \right),$$

where Φ_u is a hard-thresholding operator [54] and k denotes the iteration index. The tensor $\mathcal{R}_i$ is obtained by unfolding (tensorizing) the matrix $\mathbf{R}_{i(i)}$

Proposition 2: Given $\mathcal{R}_i^{k+1}, \mathcal{C}_i^k$ ($i = 1, \dots, N$) problem (10) has a closed form solution as follows:

$$\mathcal{L}^{k+1} = \frac{1}{N} \sum_{i=1}^N \mathcal{C}_i^k + \zeta \mathcal{R}_i.$$

Proposition 3: Given $\mathcal{L}^{k+1}, \mathcal{M}^k$, and $\mathcal{Y}^k$, problem (11) has a closed form solution as follows:

$$\mathcal{S}^{k+1} = \frac{\gamma}{\rho} \Theta_{\frac{\lambda}{\rho}} \left(\frac{\rho}{\gamma} \left(\mathcal{X} - \mathcal{M}^k + \frac{1}{\rho} \mathcal{Y}^k \right) - \left(\mathcal{L}^{k+1} - \frac{\alpha}{\gamma} \mathcal{H} \right) \right),$$

where Θ_u is a soft-thresholding operator [54].

Proposition 4: Given $\mathcal{L}^{k+1}, \mathcal{S}^{k+1}$, and $\mathcal{Y}^k$, problem (12) has a closed form solution as follows:

$$\mathcal{M}^{k+1} = \frac{1}{\beta + \rho} \left(\mathcal{Y}^k + \rho \left(\mathcal{X} - \mathcal{L}^{k+1} - \mathcal{S}^{k+1} \right) - \beta \mathcal{H} \right).$$

The complete ADMM algorithm for the proposed method is presented in Algorithm 1.

Algorithm 1 ADMM Solver for the Proposed RSTCM Model

- 1: **Input:** The observed tensor $\mathcal{X}_t$; the low-rank tensor $\mathcal{L}_{t-1}$; parameters λ , α , and β
 - 2: **Initialization:** $\mathcal{L}_t^0 = \mathcal{S}_t^0 = \mathcal{M}_t^0 = \mathcal{Y}_t^0 = 0$; $\epsilon = 1e - 5$; $\gamma = 1.2$; $\rho_0 = 1e - 3$; $\rho_{\max} = 1e8$
 - 3: **while** not converged **do**
 - 4: Update $\mathcal{L}_t^{k+1}$ by Proposition 1;
 - 5: Update $\mathcal{S}_t^{k+1}$ by Proposition 2;
 - 6: Update $\mathcal{M}_t^{k+1}$ by Proposition 3;
 - 7: Update $\mathcal{Y}_t^{k+1}$ by Eq. (13);
 - 8: Update ρ_{k+1} by $\rho_{k+1} = \min(\gamma \rho_k, \rho_{\max})$;
 - 9: Check the convergence conditions
 $\|\mathcal{L}_t^{k+1} - \mathcal{L}_t^k\|_F \leq \epsilon$; $\|\mathcal{S}_t^{k+1} - \mathcal{S}_t^k\|_F \leq \epsilon$;
 $\|\mathcal{M}_t^{k+1} - \mathcal{M}_t^k\|_F \leq \epsilon$;
 $\|\mathcal{L}_t^{k+1} + \mathcal{S}_t^{k+1} + \mathcal{M}_t^{k+1} - \mathcal{X}_t\|_F \leq \epsilon$
 - 10: **end while**
 - 11: **Output:** $\mathcal{L}_t, \mathcal{S}_t, \mathcal{M}_t$
-

B. EWMA Control Chart for Change Detection

This section presents a procedure for monitoring higher-order data streams to detect abrupt changes caused by a shock different from the underlying dynamics of data. For each tensor data $\mathcal{X}_t \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, the proposed method estimates the sparse outlier tensor $\mathcal{S}_t$. We consider $\|\mathcal{S}_t\|_1$ as the monitoring statistic and use the exponentially weighted moving average (EWMA) control chart. A control chart characterizes the usual patterns in historical data to detect unprecedented shifts in the variations of incoming data. More specifically, a control chart defines upper and lower limits (control limits) that identify a region where the observed values follow the underlying null distribution (usual patterns) with some level of confidence. When a point falls outside these thresholds, that point is considered to be significantly different from the usual patterns, indicating an abrupt change in the underlying process. A control chart can be viewed as dynamic hypothesis testing. The monitoring process using control charts consists of two phases. In Phase I, historical data that are obtained when the process is in-control and does

not contain abnormal observations (also known as in-control data) is used to determine the control limits. In Phase II, new streams of data are constantly monitored by comparing the monitoring statistics against the control limits to detect statistically significant abnormal traffic patterns. The control limits of the EWMA chart depend on the parameters of the monitoring statistic's distribution which are estimated in Phase I. Because of the definition of the monitoring statistic $\|\mathcal{S}_t\|_1 = \sum_{i_1=1}^{I_1} \dots \sum_{i_N=1}^{I_N} |(\mathcal{S}_t)_{i_1 i_2 \dots i_N}|$ and by the central limit theorem, $\|\mathcal{S}_t\|_1$ approximately follows a normal distribution. The mean and standard deviation of this normal distribution is estimated given a sequence of phase I traffic data, $\{\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_T\}$. The estimated parameters are used to obtain the control limits. More specifically, the EWMA monitoring statistic at time t is calculated by $z_t = \lambda_{cc} \|\mathcal{S}_t\|_1 + (1 - \lambda_{cc}) z_{t-1}$, where $0 < \lambda_{cc} \leq 1$ is the weighted factor of EWMA control chart. The corresponding upper control limit (UCL) and lower control limit (LCL) at time t are calculated by

$$\begin{aligned} \text{UCL}_t &= \mu_0 + L\sigma \sqrt{\frac{\lambda_{cc}}{2 - \lambda_{cc}}} [1 - (1 - \lambda_{cc})^{2t}], \\ \text{LCL}_t &= \mu_0 - L\sigma \sqrt{\frac{\lambda_{cc}}{2 - \lambda_{cc}}} [1 - (1 - \lambda_{cc})^{2t}], \end{aligned}$$

where μ_0 and σ are the in-control mean and standard deviation of $\|\mathcal{S}\|_1$ estimated in Phase I. The two parameters λ_{cc} and L are obtained through Monte Carlo simulations to achieve a desired in-control average run length (ARL_0), as detailed in the next section. The average run length (ARL) indicates how long it takes, on average, for the monitoring statistic to fall outside the control limits. Even when the process is in-control, the monitoring statistic may occasionally fall outside the control limits due to a type I error. The ARL_0 refers to the average run length when the process is in-control. Similarly, The out-of-control average run length (ARL_1) indicates how long it takes, on average, to detect a change after it occurs in a process. When in Phase II, for each incoming data $\mathcal{X}_t \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ ($t = T + 1, T + 2, \dots$), the proposed method first estimates the sparse outlier tensor $\mathcal{S}_t$ and consequently monitor the EWMA statistics z_t . If $z_t > \text{UCL}_t$ or $z_t < \text{LCL}_t$, the sample is considered as out-of-control, indicating a change has occurred in the data stream.

C. Parameter Tuning

The proposed RSTCM method involves three hyperparameters α , β , and λ . A proper setting of these parameters is essential to achieve adequate model performance. As discussed in Section III-A, the parameter λ encourages the sparsity of the outlier tensor, $\mathcal{S}$ and the larger the λ , the sparser the tensor $\mathcal{S}$ would be. The parameters α and β adjust the current estimation flexibility with respect to the previous estimations. That is, when α and β are very large (i.e., tend to infinity), the current estimation of low-rank tensor at time t becomes the same as the previous estimation at time $t - 1$. On the other hand, when α and β are set to zero, the current low-rank tensor is estimated regardless of historical data. In our implementation,

an empirical value of λ is given as $\lambda = 1/\sqrt{\max(I_1, I_2)}$ and $\lambda = 1/\sqrt{\max(I_1, I_2)I_3}$ for a two-way and a three-way tensor, respectively. According to [53], this selection of λ results in exact recovery of corrupted low-rank tensors. We select α and β by using $\frac{1}{\delta \cdot \|\mathcal{X}_{t+1}^{\Gamma_t} - \mathcal{X}_t^{\Gamma_t}\|_F}$, where $\Gamma_t = \Omega_{t+1} \cap \Omega_t$. Particularly, assuming we have T observations during Phase I (i.e., historical data), we set $\alpha = \beta = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{1}{\delta \cdot \|\mathcal{X}_{t+1}^{\Gamma_t} - \mathcal{X}_t^{\Gamma_t}\|_F}$. The intuition behind this selection is that when $\mathcal{X}_t$ is very similar to $\mathcal{X}_{t+1}$, (i.e., gradual temporal variations are very small), the values of α and β increase, which encourages the similarity between the current and previous estimations of the low-rank tensors at times t and $t+1$. On the other hand, when the observed data from time t and $t+1$ are considerably different (i.e., the data shows natural temporal variations), these parameters are set to smaller values, allowing for a more flexible estimation of the low-rank tensors. For example, if the historical data remains constant over time (i.e., no temporal patterns exist), both α and β are set to infinity, which forces the estimations of $\mathcal{L}_t$ and $\mathcal{L}_{t+1}$ be equal, representing no temporal changes in the low-rank estimations. Finally, we initiate the learning rate $\rho = 0.001$ and dynamically increase it to reach a maximum value $\rho_{max} = 10^8$. The dynamic increase of ρ ensures the constraints are satisfied and the solution is feasible.

IV. PERFORMANCE EVALUATION USING SIMULATIONS

This section examines the performance of the proposed RSTCM using synthetic datasets where the data streams contains sparse corruptions as outliers. We compare the proposed method with five benchmarks. The first benchmark, denoted by RTC, uses the robust tensor completion method introduced in Section II-B and monitors $\|\mathcal{S}\|_1$ by an EWMA control chart. The second benchmark, designated as UPCA, employs unfolded principal component analysis to extract low-dimensional features. UPCA transforms the data into a matrix by unfolding the tensor along one of the tensor modes (e.g., first mode) and then applies general principal component analysis (PCA) to the unfolded tensor to extract features. The third benchmark, denoted by MPCA, uses multilinear PCA proposed by [25] to extract features from data. The fourth benchmark, identified by TROD, extracts data features by factorizing a tensor into a sum of rank one tensors. Finally, the fifth benchmark, designated as UMPCA, extracts data features through uncorrelated multilinear principal component analysis, introduced by [26]. Features extracted from UPCA, MPCA, TROD, and UMPCA are monitored using a Hotelling T^2 control chart. When running these algorithms, we follow the settings as detailed in [62].

A. Data Generation

The data-generating process used in the experiments is defined by the first-order autoregressive (AR(1)) model. Following the AR(1) model, we generate a sequence of third-order tensors $\{\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_T\}$ in an in-control situation. Especially, at each acquisition time t , we generate a tensor data $\mathcal{X}_t = \mathcal{L} + \phi \mathcal{X}_{t-1} + \mathcal{N}_t \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, where $\phi = 0.9$ and the tensor $\mathcal{L} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is generated via three factor matrices

$\mathbf{U}^{(i)} \in \mathbb{R}^{I_i \times R}$ ($i = 1, 2, 3$) whose entries are randomly drawn from a standard normal distribution. More specifically, $\mathcal{L} = \llbracket \mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)} \rrbracket$ and $\mathcal{N}_t \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is the global Gaussian noise tensor, whose entries are simulated from a normal distribution $N(0, (k_1 \sigma_{\mathcal{L}})^2)$, where k_1 is a positive number and $\sigma_{\mathcal{L}}$ is the standard deviation of the tensor $\mathcal{L}$. Starting from time $T+1$, a sparse tensor of outliers is added to generate out-of-control samples. Specifically, the observed tensor is generated by $\mathcal{X}_t = \mathcal{L} + \phi \mathcal{X}_{t-1} + \mathcal{N}_t + \mathcal{S}$ where $\mathcal{S}$ is a sparse tensor that represents the temporal changes. The sparse tensor is generated as follows: We simulate all the entries of $\mathcal{S}$ from a normal distribution $N(0, (k_2 \sigma_{\mathcal{L}})^2)$, where k_2 is a positive scalar. Then, we randomly keep a certain proportion of entries of $\mathcal{S}$, denoted by c , and set the rest $1 - c$ percent to zero. In other words, the parameters c and k_2 control the occurrence probability of corruption and the corruption level.

At each time t , we randomly keep p percent of tensor entries $\mathcal{X}_t$ and eliminate the rest $1 - p$ percent of entries to represent missing values. Simulation studies are conducted in two stages. First, we generate 1,000 in-control sample sequences of length 1,000 to determine the EWMA control chart parameter L , while fixing λ_{cc} to 0.9, so that the in-control average run length (ARL_0) is approximately 200. Secondly, 300 out-of-control samples are generated to evaluate the change detection performance of the proposed method. The average out-of-control run length ARL_1 computed over 1000 replications is used as the evaluation criterion. For a fixed ARL_0 , a control chart with a smaller ARL_1 can detect changes more quickly.

B. Parameter Settings

In this simulation, we set the tensor size and the tensor rank as $I_1 = I_2 = I_3 = 10$, $R = 3$. The Gaussian noise level parameter is set as $k_1 = 0.1$. We also select c and k_2 from the sets $c \in \{0.1, 0.2\}$ and $k_2 \in \{0.3, 0.4, 0.5, 0.6, 0.7\}$ to validate the performance of our proposed method. The parameter L is found as $L = 2.8906$. Subsequently, we also test the effectiveness of our method on different levels of partially observed data. Specifically, we set $p \in \{1, 0.9, 0.8\}$. The parameter δ is set to 0.1. For a fair comparison, the low-rank tensor completion method is first applied to benchmarks UPCA, MPCA, UMPCA, and TROD, since they are not designed for partially observed data.

C. Simulation Results

In this subsection, we demonstrate and analyze the results from our simulation experiments. Figure 1 demonstrates the estimated ARL_1 values of different methods in Phase II for the described simulation. According to Figure 1, our proposed method outperforms other benchmarks in detecting the outliers. In other words, RSTCM exhibits the best performance in the quick detection of abrupt changes. For example, when $k_1 = 0.1$, $c = 0.2$, $p = 1$, and $k_2 = 0.5$, the average run length of each method is as follows: RSTCM: 6.1, RTC: 8.2, TROD: 21.7, MPCA: 24.3, UMPCA: 20.8, UPCA: 43.7. This indicates that the proposed robust tensor completion methods can effectively capture anomalies due to its capacity to model spatial and temporal characteristics of the data and

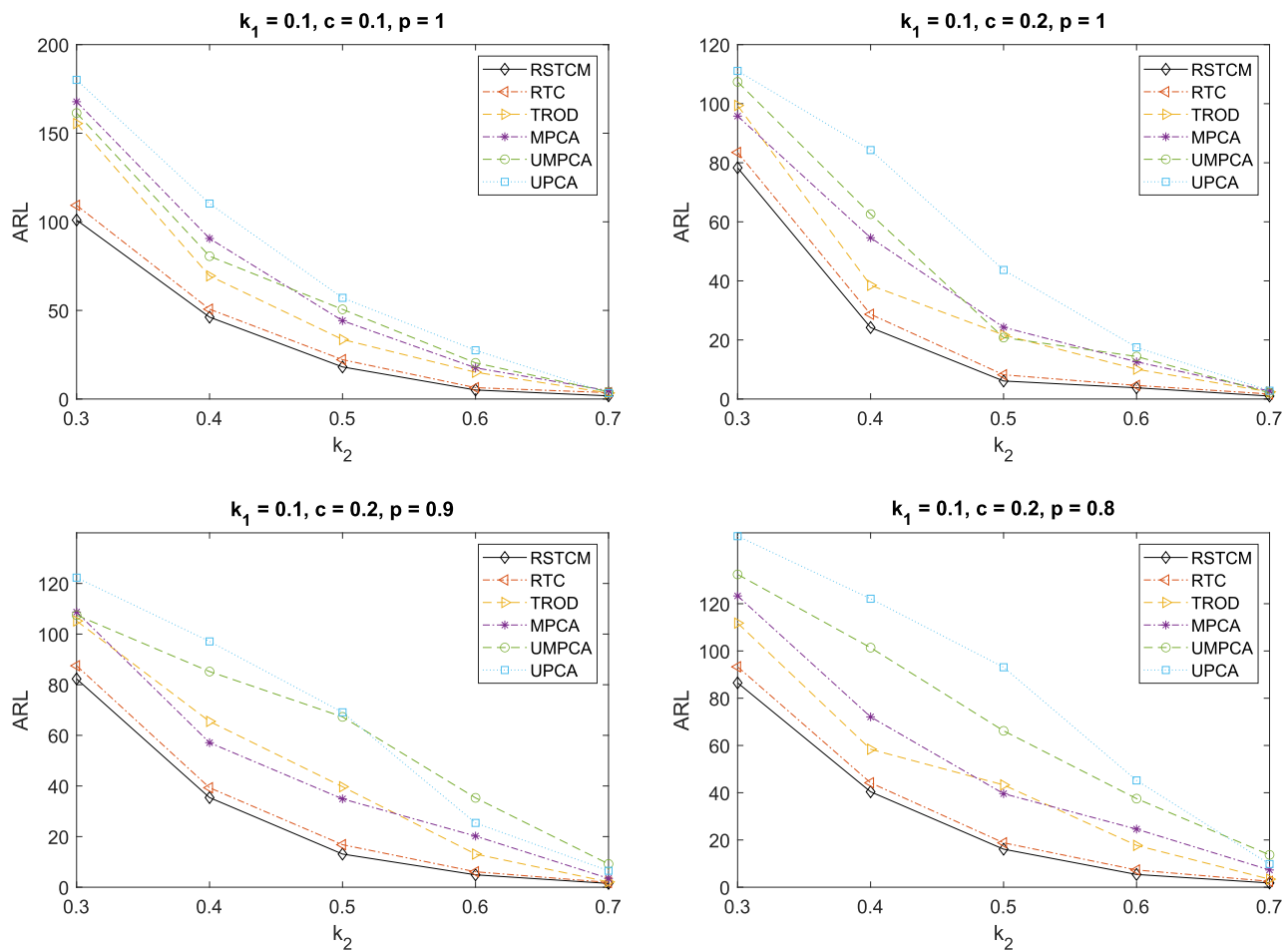


Fig. 1. Average Run Length (ARL) of the proposed method in comparison to the benchmarks.

separate the outliers. More specifically, our proposed method can detect the smaller changes (i.e., $k < 1$) faster (with a smaller standard error) than all benchmarks in all cases. For large changes (i.e., $k > 2.5$), the performance of all methods is comparable. Among all benchmarks, UPCA indicates the worst performance. Besides the fact that UPCA cannot capture the temporal pattern of data, the main reason that UPCA fails to detect small changes is that this method destroys the inner structure of tensor data.

V. IMPLEMENTATION AND EMPIRICAL EVALUATION: NEW YORK CITY TRAFFIC NETWORK DURING HURRICANE SANDY

In this section, we implement the proposed monitoring and change detection method using historical traffic data in New York City (NYC). Using this dataset, we evaluate the performance of the proposed solution in detecting unusual traffic patterns in the NYC road network caused by Hurricane Sandy in 2012. Hurricane Sandy was formed on October 22, 2012, and hit New York City on October 29, 2012. Even though it was only a Category 1 hurricane when it made landfall in the NYC area, considering the dense urban areas that were affected, the storm caused major disruptions in many infrastructure systems, including road networks.

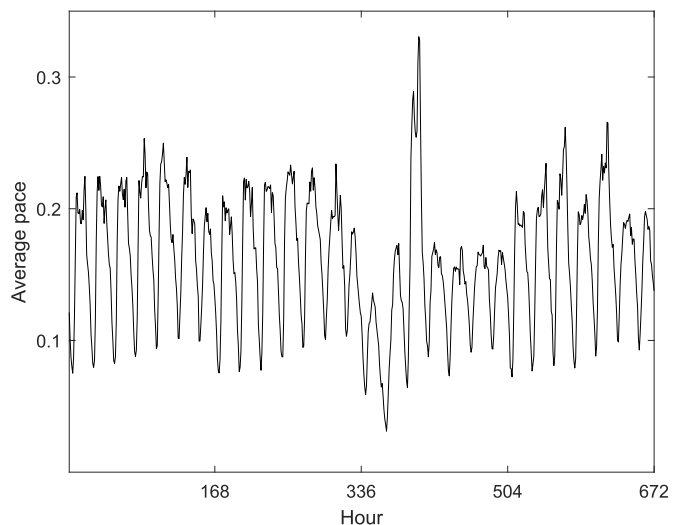


Fig. 2. Hourly average pace of traffic during the selected four-week period (October 14, 2012, to November 11, 2012).

We selected this case study to evaluate the performance of the method because of three reasons. First, the NYC road network is a good example of a complex traffic network with complex interconnectivities generating HD traffic data. Second, historical traffic data in NYC is publicly available.

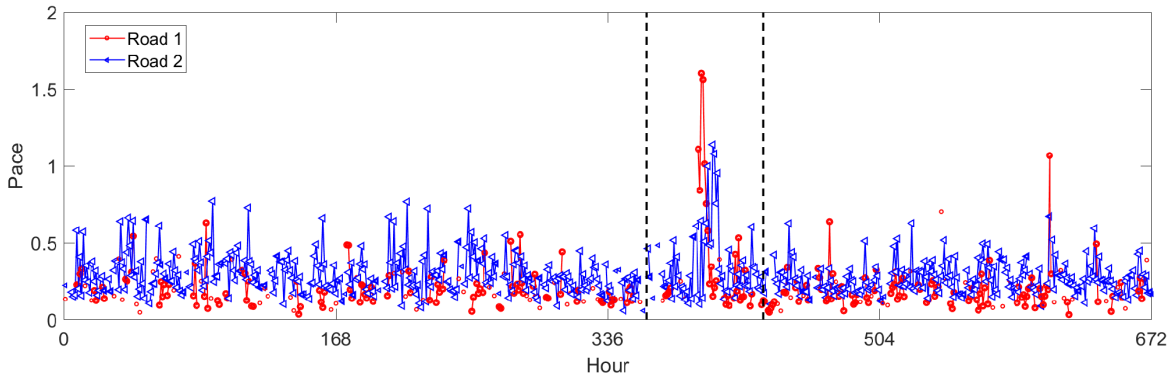


Fig. 3. Example of traffic data of two roads over 672 hours (October 14 to November 11, 2012). Data from both roads shows a significant level of missing values. The period between two black dash lines shows hurricane Sandy's duration in New York City.

Third, a series of studies have used this dataset to analyze the impacts of the hurricane on transportation networks. Their outcomes provide a benchmark to evaluate and validate the results of our assessment.

The NYC historical traffic dataset, prepared by Donovan and Work [63], contains hourly average travel time on each road segment of the network. The average travel times were estimated using the historical trajectory of nearly 700 million taxi trips collected from more than 100,000 taxis equipped with GPS trackers in NYC from January 2010 to December 2013. The NYC road network in the borough of Manhattan consists of 8839 links (i.e., road segments) and 3910 nodes. The dataset also includes the length of each road segment, which allows us to calculate the hourly average speed and pace (i.e., the inverse of speed) on each road segment.

As suggested by previous studies [17], [63], traffic conditions in each road of the network are quantified using traffic pace, which is calculated as travel time divided by the length of the road segment. In this study, a subset of the data, including a period of four weeks from October 14, 2012 (two weeks before the hurricane) to November 11, 2012 (two weeks after the hurricane), is used to analyze the impacts of the hurricane. This subset of data forms a stream of 672 tensors representing the traffic conditions for all the 8839 road segments during the four-week time window. The daily average pace in a normal day before and after the hurricane is around 0.16, however, the daily average pace during the hurricane week is 0.13 (see the figure 2). This dataset contains 32% missing observations. Figure 3 shows the observed and missing data in two road segments as an example. In this figure, the black dashes indicate the approximate time of the hurricane, and the unconnected lines (i.e., blue and red) denote that there are missing values in between. This figure shows that the percentage of missing observations considerably increased when the hurricane hit the city.

Considering cyclical patterns in traffic pace during every 24 hour, we consider a sliding window of size 8839×24 with a sliding step of $s = 4$ that moves across the complete matrix of traffic data of size 8839×672 . Therefore, 163 matrices of size 8839×24 are finally obtained as samples to be monitored. Next, the first week's data, which contains 43 matrices, are considered in-control samples. In-control samples are assumed

TABLE I

THE NUMBER OF HOURS BEFORE (-) OR AFTER (+) THE HURRICANE LANDFALL EACH METHOD DETECTS UNUSUAL TRAFFIC PATTERNS

	UPCA	MPCA	UMPCA	TROD	RSTCM
Hours	+88h	+16h	-3h	+56h	-15h

to show usual traffic patterns and are used to determine the control limits for the EWMA control chart (known as phase I). The remaining 120 samples may contain unusual and anomalous traffic patterns and are tested for change detection in the second phase of the statistical process control. The parameters of the EWMA control chart are set as $L = 3$ and $\lambda_{cc} = 1$ and δ is set as 0.1. The EWMA control chart and the variations of the monitoring statistics (i.e., $\|\mathcal{S}\|_1$) obtained by our method are plotted in Figure 4.

The results indicate traffic patterns in the NYC network experienced abnormal changes from sample 84 (i.e., October 29, 2012, at 9:00 PM) to sample 112 (i.e., November 3, 2012, at 12:00 PM). These findings are closely aligned with the results from previous studies [16], [17], confirming the validity and accuracy of our method in detecting disruptions at the network level. It is worth noting that the previous studies aggregate the data at the network level to handle missing values and cannot identify the local disruptions. Our proposed method directly accounts for missing values and allows for fine-granularity analysis of the network. To evaluate the performance of the proposed methods in the monitoring of the traffic data at the network level, we compare it with the benchmark methods (UPCA, MPCA, UMPCA and TROD) presented in Section IV. Considering that these methods cannot handle missing observations, we first apply low-rank tensor completion to impute the missing values. Table I reports the number of hours from the time of hurricane landfall (used as a reference point) that each method detects network-level disruptions. The results indicate that the proposed method outperforms other benchmarks and can detect unusual traffic patterns around 15 hours before the hurricane's landfall.

Another major advantage of the proposed method in this study is the ability to detect unusual local traffic changes at the road segment level of granularity. Figure 5 shows the NYC road network one day after the hurricane's landfall. The road segments highlighted in red were experiencing abnormal

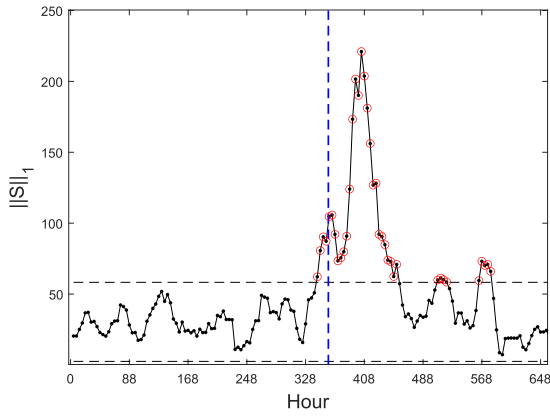


Fig. 4. The monitoring statistic based on the traffic data in New York City (The blue line shows when hurricane Sandy landfalls).



Fig. 5. Gridlocks in the NYC road network one day after the Hurricane Sandy landfall.

gridlocks. It should be noted that many roads during that period of time did not carry a significant amount of traffic since the hurricane disrupted almost all daily operations and most citizens sheltered at their homes or evacuated before the hurricane hit the city. However, some road segments, such as those closer to the entrance of the Lincoln Tunnel and the bridge to Roosevelt Island, experienced unusually heavy traffic. The performance of the proposed method in detecting local disruptions at the road segment level can be evaluated by visually inspecting the time series of traffic pace in the road segments. For example, the results of the proposed monitoring method indicated that road segment 33362, which is part of 8th Avenue near the entrance of the Lincoln Tunnel, experienced an abnormal gridlock after the hurricane's landfall. Figure 6 (a) shows the hourly traffic pace in this road segment during the four-week analysis period, validating the significant shift in traffic pace variations. Figure 6 (b) shows the time series of traffic pace in road segment 97082. This road segment is part of 3rd Avenue in residential areas of the Upper East Side. The time series does not show any significant shift in the

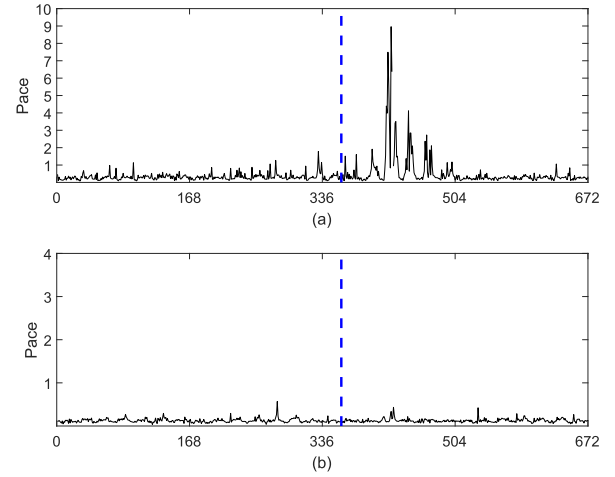


Fig. 6. Hourly traffic pace on two sample road segments with different levels of disruptions during Hurricane Sandy. The blue dashed line denotes the start of Hurricane Sandy.

variations of the traffic pace during the analysis period. This observation is closely aligned with the results of the proposed method that did not detect any traffic gridlock in this road segment.

VI. CONCLUSION

The primary contribution of this study to the core body of knowledge is to develop and empirically evaluate a traffic monitoring solution for incomplete and partially observed HD traffic data using a robust spatiotemporal tensor completion-based monitoring (RSTCM) method. The proposed method exploits the spatial and temporal patterns of the data to complete and decompose the tensor data in the sum of low-rank and sparse tensors. The sparse tensor represents the outliers and is used to define monitoring statistics to be monitored using an EWMA control chart. We examined the performance of the proposed method using two methods: (1) a simulation analysis using synthetic data and (2) an empirical assessment using historical traffic data in NYC to assess the method's performance in detecting congestions and unusual traffic patterns caused by Hurricane Sandy in 2012. The outcomes of both simulation and empirical assessments confirmed the accuracy and efficacy of the proposed method. The proposed method uses a sliding window of time to create the tensor of traffic data and assumes an autoregressive underlying model, which may limit the memory of the algorithm in capturing historical patterns. To address these limitations, future studies can extend the proposed method by incorporating deep neural network methods, such as long short-term memory (LSTM), to capture both short and long-term patterns in the underlying dynamics of the data. The LSTM will serve as a constraint on the smooth part of the model.

ACKNOWLEDGMENT

Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- [1] H. Yan, K. Paynabar, and J. Shi, "Real-time monitoring of high-dimensional functional data streams via spatio-temporal smooth sparse decomposition," *Technometrics*, vol. 60, no. 2, pp. 181–197, Apr. 2018.

- [2] M. Nouri, M. Reisi-Gahrooei, and M. Ilbeigi, "A method for granular traffic data imputation based on PARATUCK2," *Transp. Res. Rec.*, vol. 2676, no. 10, pp. 220–230, Oct. 2022.
- [3] Y. Hu and D. B. Work, "Robust tensor recovery with fiber outliers for traffic events," *ACM Trans. Knowl. Discovery Data*, vol. 15, no. 1, pp. 1–27, Feb. 2021.
- [4] J. Evans, B. Waterson, and A. Hamilton, "A random forest incident detection algorithm that incorporates contexts," *Int. J. Intell. Transp. Syst. Res.*, vol. 18, no. 2, pp. 230–242, May 2020.
- [5] H. W. Li, S. L. Li, H. W. Zhu, X. Zhao, and X. Zhang, "Automated detection algorithm for traffic incident in urban expressway based on lengthways time series," in *Proc. 8th Int. Conf. Green Intell. Transp. Syst. Saf.* Singapore: Springer, 2019, pp. 625–633.
- [6] J. E. Abanto, C. C. Reyes, J. A. Malinao, and H. N. Adorna, "Traffic incident detection and modelling using quantum frequency algorithm and AutoRegressive integrated moving average models," in *Proc. IISA*, Jul. 2013, pp. 1–6.
- [7] S. Tang and H. Gao, "Traffic-incident detection-algorithm based on nonparametric regression," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 1, pp. 38–42, Mar. 2005.
- [8] L. Li, Y. Lin, B. Du, F. Yang, and B. Ran, "Real-time traffic incident detection based on a hybrid deep learning model," *Transportmetrica A, Transp. Sci.*, vol. 18, no. 1, pp. 78–98, Mar. 2022.
- [9] L. Zhu, F. Guo, R. Krishnan, and J. W. Polak, "A deep learning approach for traffic incident detection in urban networks," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, 2018, pp. 1011–1016.
- [10] D. Srinivasan, X. Jin, and R. L. Cheu, "Adaptive neural network models for automatic incident detection on freeways," *Neurocomputing*, vol. 64, pp. 473–496, Mar. 2005.
- [11] J. Xiao, "SVM and KNN ensemble learning for traffic incident detection," *Phys. A, Stat. Mech. Appl.*, vol. 517, pp. 29–35, Mar. 2019.
- [12] A. B. Parsa, H. Taghipour, S. Derrible, and A. Mohammadian, "Real-time accident detection: Coping with imbalanced data," *Accident Anal. Prevention*, vol. 129, pp. 202–210, Aug. 2019.
- [13] B. Yao, P. Hu, M. Zhang, and M. Jin, "A support vector machine with the Tabu search algorithm for freeway incident detection," *Int. J. Appl. Math. Comput. Sci.*, vol. 24, no. 2, pp. 397–404, Jun. 2014.
- [14] Y. Sun, T. Mallick, P. Balaprakash, and J. Macfarlane, "A data-centric weak supervised learning for highway traffic incident detection," *Accident Anal. Prevention*, vol. 176, Oct. 2022, Art. no. 106779.
- [15] N. Dogru and A. Subasi, "Traffic accident detection using random forest classifier," in *Proc. 15th Learn. Technol. Conf. (LT)*, 2018, pp. 40–45.
- [16] B. Donovan and D. B. Work, "Empirically quantifying city-scale transportation system resilience to extreme events," *Transp. Res. C, Emerg. Technol.*, vol. 79, pp. 333–346, Jun. 2017.
- [17] M. Ilbeigi, "Statistical process control for analyzing resilience of transportation networks," *Int. J. Disaster Risk Reduction*, vol. 33, pp. 155–161, Feb. 2019.
- [18] C. Zou and P. Qiu, "Multivariate statistical process control using LASSO," *J. Amer. Stat. Assoc.*, vol. 104, no. 488, pp. 1586–1596, Dec. 2009.
- [19] C. Zou, X. Ning, and F. Tsung, "LASSO-based multivariate linear profile monitoring," *Ann. Operations Res.*, vol. 192, no. 1, pp. 3–19, Jan. 2012.
- [20] Y. Zhao, H. Yan, S. Holte, and Y. Mei, "Rapid detection of hot-spots via tensor decomposition with applications to crime rate data," *J. Appl. Statist.*, vol. 49, no. 7, pp. 1636–1662, May 2022.
- [21] B. R. Bakshi, "Multiscale PCA with application to multivariate statistical process monitoring," *AIChE J.*, vol. 44, no. 7, pp. 1596–1610, Jul. 1998.
- [22] B. De Ketelaere, M. Hubert, and E. Schmitt, "Overview of PCA-based statistical process-monitoring methods for time-dependent, high-dimensional data," *J. Quality Technol.*, vol. 47, no. 4, pp. 318–335, Oct. 2015.
- [23] R. Y. Liu, "Control charts for multivariate processes," *J. Amer. Stat. Assoc.*, vol. 90, no. 432, pp. 1380–1387, Dec. 1995.
- [24] K. Paynabar, C. Zou, and P. Qiu, "A change-point approach for phase-I analysis in multivariate profile monitoring and diagnosis," *Technometrics*, vol. 58, no. 2, pp. 191–204, Apr. 2016.
- [25] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "MPCA: Multilinear principal component analysis of tensor objects," *IEEE Trans. Neural Netw.*, vol. 19, no. 1, pp. 18–39, Jan. 2008.
- [26] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Uncorrelated multilinear principal component analysis through successive variance maximization," in *Proc. 25th Int. Conf. Mach. Learn.*, 2008, pp. 616–623.
- [27] D. Qian et al., "Bayesian nonnegative CP decomposition-based feature extraction algorithm for drowsiness detection," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 8, pp. 1297–1308, Aug. 2017.
- [28] Z. Hou, W. Li, R. Tao, and Q. Du, "Three-order Tucker decomposition and reconstruction detector for unsupervised hyperspectral change detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 6194–6205, 2021.
- [29] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM (JACM)*, vol. 58, no. 3, pp. 1–37, 2011.
- [30] C. Schwartz et al., "Change detection in UWB SAR images based on robust principal component analysis," *Remote Sens.*, vol. 12, no. 12, p. 1916, Jun. 2020.
- [31] L. P. Ramos et al., "A wavelength-resolution SAR change detection method based on image stack through robust principal component analysis," *Remote Sens.*, vol. 13, no. 5, p. 833, Feb. 2021.
- [32] H. Zhang, Q. Huang, H. Zhai, and L. Zhang, "Multi-temporal cloud detection based on robust PCA for optical remote sensing imagery," *Comput. Electron. Agric.*, vol. 188, Sep. 2021, Art. no. 106342.
- [33] B. Park, C. J. Messer, and T. Urbanik, "Short-term freeway traffic volume forecasting using radial basis function neural network," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1651, no. 1, pp. 39–47, Jan. 1998.
- [34] M. Zhong, S. Sharma, and P. Lingras, "Genetically designed models for accurate imputation of missing traffic counts," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1879, no. 1, pp. 71–79, Jan. 2004.
- [35] C. Zhang, S. Sun, and G. Yu, "A Bayesian network approach to time series forecasting of short-term traffic flows," in *Proc. 7th Int. IEEE Conf. Intell. Transp. Syst.*, Mar. 2004, pp. 216–221.
- [36] B. Ghosh, B. Basu, and M. O'Mahony, "Bayesian time-series model for short-term traffic flow forecasting," *J. Transp. Eng.*, vol. 133, no. 3, pp. 180–189, Mar. 2007.
- [37] H. Dia, "An object-oriented neural network approach to short-term traffic forecasting," *Eur. J. Oper. Res.*, vol. 131, no. 2, pp. 253–261, Jun. 2001.
- [38] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach," *Transp. Res. C, Emerg. Technol.*, vol. 13, no. 3, pp. 211–234, Jun. 2005.
- [39] M. Castro-Neto, Y.-S. Jeong, M.-K. Jeong, and L. D. Han, "Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions," *Exp. Syst. Appl.*, vol. 36, no. 3, pp. 6164–6173, Apr. 2009.
- [40] X. Jin, Y. Zhang, and D. Yao, "Simultaneously prediction of network traffic flow based on PCA-SVR," in *Proc. Int. Symp. Neural Netw.* Berlin, Germany: Springer, Jun. 2007, pp. 1022–1031.
- [41] W. Yin, P. Murray-Tuite, and H. Rakha, "Imputing erroneous data of single-station loop detectors for nonincident conditions: Comparison between temporal and spatial methods," *J. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 159–176, Jul. 2012.
- [42] M. Zhong, S. Sharma, and Z. Liu, "Assessing robustness of imputation models based on data from different jurisdictions: Examples of Alberta and Saskatchewan, Canada," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1917, pp. 116–126, Jan. 2005.
- [43] O. Troyanskaya et al., "Missing value estimation methods for DNA microarrays," *Bioinformatics*, vol. 17, no. 6, pp. 520–525, Jun. 2001.
- [44] D. Ni and J. D. Leonard, "Markov chain Monte Carlo multiple imputation using Bayesian networks for incomplete intelligent transportation systems data," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1935, no. 1, pp. 57–67, Jan. 2005.
- [45] L. Qu, L. Li, Y. Zhang, and J. Hu, "PPCA-based missing data imputation for traffic flow volume: A systematical approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 3, pp. 512–522, Sep. 2009.
- [46] Y. Li, Z. Li, and L. Li, "Missing traffic data: Comparison of imputation methods," *IET Intell. Transp. Syst.*, vol. 8, no. 1, pp. 51–57, Feb. 2014.
- [47] X. Chen, Z. Wei, Z. Li, J. Liang, Y. Cai, and B. Zhang, "Ensemble correlation-based low-rank matrix completion with applications to traffic data imputation," *Knowl.-Based Syst.*, vol. 132, pp. 249–262, Sep. 2017.
- [48] R. Du, C. Chen, B. Yang, and X. Guan, "VANET based traffic estimation: A matrix completion approach," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2013, pp. 30–35.
- [49] D. Zhang, Y. Hu, J. Ye, X. Li, and X. He, "Matrix completion by truncated nuclear norm regularization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2192–2199.

- [50] R. Ma, N. Barzigar, A. Roozgard, and S. Cheng, "Decomposition approach for low-rank matrix completion and its applications," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1671–1683, Apr. 2014.
- [51] P. Jain, P. Netrapalli, and S. Sanghavi, "Low-rank matrix completion using alternating minimization," in *Proc. 45th Annu. ACM Symp. Theory Comput.*, Jun. 2013, pp. 665–674.
- [52] D. Goldfarb and Z. Qin, "Robust low-rank tensor recovery: Models and algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 35, no. 1, pp. 225–253, Jan. 2014, doi: [10.1137/130905010](https://doi.org/10.1137/130905010).
- [53] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 5249–5257.
- [54] N. Xue, G. Papamakarios, M. Bahri, Y. Panagakis, and S. Zafeiriou, "Robust low-rank tensor modelling using Tucker and CP decomposition," in *Proc. 25th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2017, pp. 1185–1189.
- [55] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.
- [56] X. Zhao, M. Bai, D. Sun, and L. Zheng, "Robust tensor completion: Equivalent surrogates, error bounds, and algorithms," *SIAM J. Imag. Sci.*, vol. 15, no. 2, pp. 625–669, Jun. 2022.
- [57] D. Qiu, M. Bai, M. K. Ng, and X. Zhang, "Robust low-rank tensor completion via transformed tensor nuclear norm with total variation regularization," *Neurocomputing*, vol. 435, pp. 197–215, May 2021.
- [58] G. Song, M. K. Ng, and X. Zhang, "Robust tensor completion using transformed tensor singular value decomposition," *Numer. Linear Algebra Appl.*, vol. 27, no. 3, p. e2299, May 2020.
- [59] Q. Jiang and M. Ng, "Robust low-tubal-rank tensor completion via convex optimization," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, Aug. 2019, pp. 2649–2655.
- [60] M. Zhao, M. Reisi Gahrooei, and N. Gaw, "Robust coupled tensor decomposition and feature extraction for multimodal medical data," *IJSE Trans. Healthcare Syst. Eng.*, vol. 13, no. 2, pp. 117–131, Apr. 2023.
- [61] F. Wang, M. R. Gahrooei, Z. Zhong, T. Tang, and J. Shi, "An augmented regression model for tensors with missing values," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 4, pp. 2968–2984, Oct. 2022.
- [62] H. Yan, K. Paynabar, and J. Shi, "Image-based process monitoring using low-rank tensor decomposition," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 1, pp. 216–227, Jan. 2015.
- [63] B. Donovan and D. B. Work, "New York City taxi trip data (2010–2013)," Dataset, 2014, doi: [10.13012/J8PN93H8](https://doi.org/10.13012/J8PN93H8).



Meng Zhao received the Ph.D. degree from the Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL, USA, in 2023. She is currently a Post-Doctoral Research Scientist with Columbia University. Her research interests include high-dimensional data modeling, tensor analysis, and machine learning techniques. She is a member of the Institute for Operations Research and the Management Science (INFORMS) and the Institute of Industrial and Systems Engineers (IISE).



Mostafa Reisi Gahrooei received the master's degree in computational science and engineering and the Ph.D. degree in industrial and systems engineering from Georgia Institute of Technology, Atlanta, GA, USA, and the M.Sc. degree in transportation engineering and applied mathematics from Southern Illinois University Edwardsville, Edwardsville, IL, USA. He is currently an Assistant Professor with the Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL, USA. His research interests include modeling and monitoring complex systems with functional and high-dimensional data. He is a member of the Institute for Operations Research and the Management Science (INFORMS) and the Institute of Industrial and Systems Engineers (IISE).



Mohammad Ilbeigi received the B.Sc. degree in civil engineering from the Sharif University of Technology, Iran, the M.Sc. degree in civil engineering from the University of Tehran, Iran, and the M.Sc. degree in computational science and engineering and the Ph.D. degree in construction management from Georgia Institute of Technology, Atlanta, GA, USA. He is currently an Assistant Professor with the Department of Civil, Environmental, and Ocean Engineering, Stevens Institute of Technology, Hoboken, NJ, USA. His research interests include resilient and adaptive infrastructure development.