

Detecting Traffic Anomalies During Extreme Events via a Temporal Self-Expressive Model

Mina Nouri^{ID}, Elif Konyar^{ID}, Mostafa Reisi Gahrooeri^{ID}, and Mohammad Ilbeigi^{ID}

Abstract—Motivated by rapid urbanization and increasing natural hazards, this study aims to develop a data-driven method for detecting urban traffic anomalies during extreme events. Past experiences have shown that abnormal traffic patterns caused by extreme events can disrupt traffic in a large portion of the road network. Timely and reliable traffic monitoring for detection of such anomalies is crucial for congestion mitigation and successful emergency operation plans. An effective traffic monitoring system should detect disruptions at both network and local levels. However, the existing methods are not capable of addressing this need. This study proposes a temporal self-expressive network monitoring method to achieve this purpose. The proposed method first utilizes a temporal self-expressive model to uncover the dynamic interdependencies between local zones of the traffic network. Next, a statistical monitoring method detects network-wide anomalies based on regular traffic interdependencies. Finally, the method identifies the zones most affected by the anomalous event. We applied the proposed method to the road network of Manhattan in New York City to evaluate its performance during Hurricane Sandy. The outcomes confirmed that the temporal self-expressive model, augmented with statistical monitoring tools, could accurately detect anomalous traffic at both network and zone levels.

Index Terms—Anomaly detection, road traffic networks, self-expressive modeling, statistical process control, urban traffic monitoring.

I. INTRODUCTION

NATURAL and man-made extreme events can significantly disrupt traffic flows on urban road networks. When not appropriately handled, traffic disruptions may lead to severe traffic congestion and unexpected gridlock, which may seriously interrupt emergency management activities, such as evacuation and recovery efforts. Past events have demonstrated the impacts of such disruptions. For example, the evacuation of Houston, Texas, before Hurricane Rita created a nearly

100-mile-long traffic gridlock that lasted an entire day and led to several deaths before the storm arrived [1], [2]. In traffic networks, extreme events can cause traffic anomalies that simultaneously affect large regions of the network. Monitoring of road traffic to detect such anomalies creates an opportunity to mitigate congestion and supports emergency operation plans.

A considerable number of studies in recent years aimed to contribute to this research topic from various perspectives. A group of studies focuses on real-time data collection to provide reliable input data for disruption detection methods. These studies consider different mediums for traffic data collection, including traffic surveillance cameras (e.g., [3], [4], [5]), loop detectors (e.g., [6], [7]), taxis and GPS trackers (e.g., [8], [9], [10]), geo-tagged data from social media (e.g., [11], [12]), and connected vehicles (e.g., [13], [14], [15]). For example, Ahmad et al. [13] proposed a traffic monitoring framework using connected vehicles (CV) trajectory data to assess the impacts of an emergency evacuation operation due to wildfire on overall traffic conditions.

The remaining studies related to this research area propose traffic disruption detection methods. These studies can be categorized into two main groups. The first group offers methods that are applicable to only a limited number of road segments. A set of these studies introduces few-step-ahead prediction methods for disruption detection. The overall approach in these studies is to use non-anomalous historical data to provide short-term forecasts of traffic flow and detect traffic states that fall outside the predictions' confidence intervals as disruptions. For example, Evans et al. [16] introduced a disruption detection algorithm that uses a context-based random forest model to forecast the traffic flow data of individual loop detectors. Their algorithm then employs the Quantile Random Tree Regression (QRTR) method to establish a prediction interval. If the observed traffic values fall outside the prediction interval three times in a row, disruption alerts are raised. Other examples include [17], [18], [19]. These methods are impractical for monitoring large-scale networks as they quickly lose their statistical power when the number of road segments increases, which results in a significant number of false alarms. The second set of studies in this group develops classification models that can differentiate normal and abnormal traffic patterns. Various machine learning methods, including Support Vector Machines (e.g., [20], [21], [22]) Random Forests (e.g., [23], [24]), and Artificial Neural Networks (e.g., [25], [26], [27]) have been used commonly in these studies. These methods are often focused on incident detection within a single roadway

Manuscript received 19 May 2023; revised 19 December 2023; accepted 21 April 2024. Date of publication 17 May 2024; date of current version 4 October 2024. This work was supported by the National Science Foundation under Grant CMMI-2027025 and Grant CMMI-2026795. The Associate Editor for this article was R. Arghandeh. (Corresponding author: Mohammad Ilbeigi.)

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

Elif Konyar and Mostafa Reisi Gahrooeri are with the Department of Industrial and Systems Engineering (ISE), University of Florida, Gainesville, FL 32603 USA (e-mail: elif.konyar@ufl.edu; mreisigahrooeri@ufl.edu).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TITS.2024.3397034>, provided by the authors.

Digital Object Identifier 10.1109/TITS.2024.3397034

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.

TABLE I
SUMMARY OF TRAFFIC ANOMALY DETECTION METHODS

Disruption Detection Methods		Literature	Characteristics
Road-level event detection methods	Prediction-based	[16], [17], [18], [19]	- Providing a short-term prediction interval for regular traffic - Not scalable to monitor large-scale road networks
	Classification-based	[20], [21], [22], [23], [24], [25], [26], [27]	- A trained model capable of differentiating normal and abnormal traffic states - Not scalable to monitor large-scale road networks
Network-level event detection methods		[8], [28], [29]	- Appropriate for detecting network-wide disruptions - Not able to localize disruptions at the zone level
Decomposition-based methods		[30], [31], [32]	- Utilizing spatiotemporal correlation of traffic data - Retroactive in nature - Not appropriate for online traffic monitoring at the network and zone levels

(e.g., a highway) and cannot be applied to a network with a large number of roads due to the increased false alarm rate.

The second group of studies proposing traffic disruption detection methods focuses on the entire road network and provides solutions to detect the presence of network-wide disruptions. One example is the study by Donovan and Work [8]. They use GPS data collected from nearly 700 million taxi trips in New York City to create probability distributions of typical travel paces and use it to identify unusual traffic patterns. However, their proposed method is computationally expensive. Therefore, they divide the entire network into four regions and develop the distributions for 16 trip types from and to each one of these regions. To address this limitation, Ilbeigi [28] introduced a computationally efficient method to detect unusual traffic patterns by monitoring traffic network topological features (e.g., closeness centrality index) using statistical process control methods such as Cumulative Sum (CUSUM) control charts. Hu et al. [29] recently introduced an autoencoder-based anomaly detection method to detect extreme events in origin-destination (OD) mobility datasets. They model the OD mobility data as a set of time-dependent directed graphs and use embedding techniques to capture the spatiotemporal patterns of normal traffic. They then detect anomalous graphs corresponding to extreme events through a semi-supervised autoencoder approach. The primary limitation of the studies in the second group contrasts with that of the first group. These studies exclusively detect large-scale events by monitoring traffic conditions at a high-level network view. However, they cannot locate specific roads (or zones) that experience abnormal traffic during such events.

An effective traffic disruption detection method should ideally monitor the entire network and its road segments concurrently. The critical element to creating such a method is to model spatiotemporal interdependencies among road segments in a network. A few studies have utilized the spatiotemporal correlation of traffic data to detect anomalous events. For example, Yang et al. [30] proposed a method for traffic event detection based on the Bayesian Robust Principal Component Analysis (BRPCA) approach. Their proposed method combines multiple traffic data streams (e.g., traffic flow and occupancy) through sharing the same sparse structure.

This sparse structure is used to localize traffic events in space and time. Li et al. [31] introduced a coupled scalable Bayesian robust tensor factorization (coupled SBRTF) model to detect non-recurrent traffic congestion (NRTC) using high-dimensional spatiotemporal patterns of traffic data. In another study, Hu and Work [32] proposed a method based on the robust tensor decomposition approach to detect large traffic events. Their proposed method separates traffic during large events from normal traffic by decomposing traffic tensor data into a low-rank tensor (for regular traffic) and a sparse outlier tensor (for large events). While these methods take advantage of the spatiotemporal correlations in traffic patterns, they come with their own limitations. These methods perform offline analysis of traffic data and detect abnormalities in a retroactive manner. This limitation makes them unsuitable for online monitoring (i.e., continuously receiving and analyzing traffic data streams to detect abnormalities) of road traffic at network and local levels of granularity.

In summary, current traffic anomaly detection methods have three main limitations: First, they are unable to simultaneously detect extreme traffic events and pinpoint specific roads (or zones) experiencing significant anomalous traffic patterns. Second, they do not fully leverage the spatiotemporal interdependencies of the traffic network to identify anomalous events. Third, many of these methods detect traffic anomalies retroactively and are not suitable for online monitoring of road networks. Table I provides a summary of the existing traffic anomaly detection methods and their characteristics.

Extreme events can significantly impact an entire road network. Therefore, it is important to develop a method that can assess the overall network traffic condition. Simultaneously, for effective congestion mitigation, this method should be capable of identifying specific regions that are most affected by such events.

Considering the limitations of existing studies, this paper proposes an innovative multi-level method for traffic network monitoring and anomaly detection. This method detects anomalies at both network and zone levels of granularity. At the network level, the proposed method models spatiotemporal interdependencies among local traffic zones to capture the network traffic condition as a whole and detect extreme

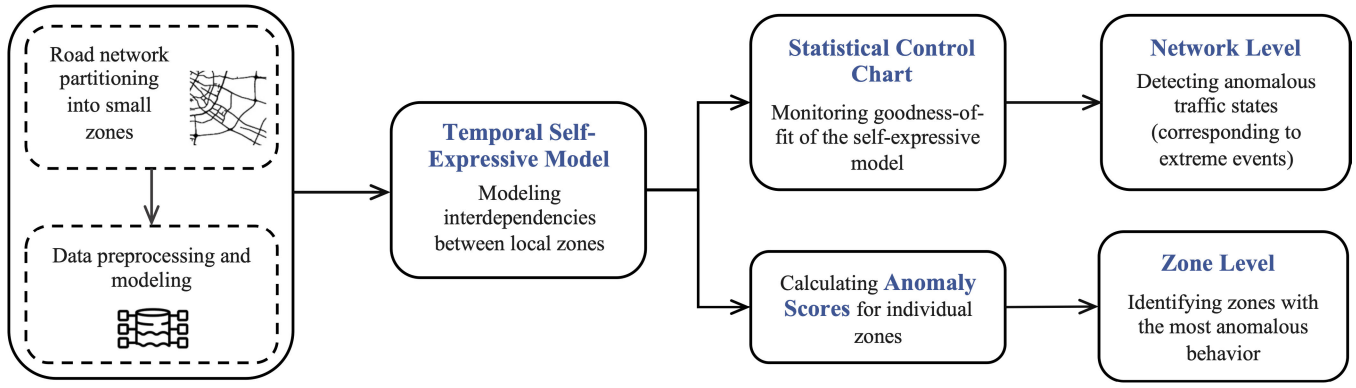


Fig. 1. The flowchart of the proposed temporal self-expressive network monitoring method.

traffic events. At the zone level, the method determines local zones that exhibit abnormal traffic behavior. In the context of extreme events, anomalies can result from either increased congestion or a substantial decrease in travel demand within a specific zone. Therefore, in this study, we define anomalous zones as those that experience remarkable deviations from their expected traffic patterns as a result of an event.

Fig. 1 provides an overview of our proposed method. In this method, first, the urban road network is divided into small zones. Next, a temporal self-expressive model is developed to characterize the dynamic traffic interdependencies between these zones. Then, using a statistical control chart, anomalous traffic states are detected at the network level through monitoring a novel goodness-of-fit measure derived from the self-expressive model. Additionally, anomaly scores are defined and calculated for individual zones over time to identify the local zones most affected by the anomalous event.

The remainder of this paper is organized as follows. In Section II, we introduce our notation. In Section III, we present the proposed methodology in detail. In Section IV, we apply the proposed method to the road network of Manhattan in New York City to detect traffic anomalies caused by Hurricane Sandy at network and zone levels. Finally, in Section V, we summarize the primary contributions of the paper to the body of knowledge and outline potential future research directions.

II. NOTATIONS AND PRELIMINARIES

The notations used in this paper follow the notations adopted by Kolda and Bader [33]. Matrices are denoted by boldface uppercase letters, e.g., $\mathbf{A}$, vectors by boldface lowercase letters, e.g., $\mathbf{a}$, and scalars by lowercase letters, e.g., a . Given a matrix $\mathbf{A} \in \mathbb{R}^I \times J$, $\mathbf{a}^i$ and $\mathbf{a}_j$ are used to represent the i^{th} row and the j^{th} column of $\mathbf{A}$, respectively. The transpose of a matrix $\mathbf{A} \in \mathbb{R}^I \times J$ is represented by $\mathbf{A}^T$, and its inverse is denoted by $\mathbf{A}^{-1}$.

We denote the l_2 norm of a vector by $\|\cdot\|_2$, which is defined as $\|\mathbf{a}\|_2 = \sqrt{\sum_k |a_k|^2}$, and the inner product of two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^K$ as $\langle \mathbf{a}, \mathbf{b} \rangle = \sum_k a_k b_k$. Moreover, $\|\mathbf{A}\|_F^2$ is used to denote the squared Frobenius norm of a given matrix $\mathbf{A} \in \mathbb{R}^I \times J$, which can be computed as the sum of the

squared l_2 norms of its column vectors. The matrix l_1 norm is represented by $\|\cdot\|_1$ and defined as $\|\mathbf{A}\|_1 = \sqrt{\sum_i \sum_j |a_{ij}|}$. Furthermore, the rank- R decomposition, also known as rank factorization, of a matrix $\mathbf{A} \in \mathbb{R}^I \times J$ has the form $\mathbf{A} = \mathbf{B}\mathbf{C}$, where $\mathbf{B} \in \mathbb{R}^I \times R$, $\mathbf{C} \in \mathbb{R}^R \times J$, and R is the rank of $\mathbf{A}$.

In addition, $\text{sgn}(\cdot)$ is the sign function, where $\text{sgn}(a) = a/|a|$ if $a \neq 0$ and $\text{sgn}(a) = 0$ if $a = 0$ for a real number a . Lastly, the notation $(a)_+$ represents the positive part of a real number a and is defined as $(a)_+ = \max\{a, 0\}$.

III. METHODOLOGY

In this section, we first present the temporal self-expressive model for characterizing road interdependencies. Next, we introduce a novel statistic that is monitored over time using statistical control charts for detecting traffic anomalies. Finally, we calculate an anomaly score for different network zones to identify regions exhibiting the most anomalous behavior.

A. Temporal Self-Expressive Modeling

The proposed temporal self-expressive model in this paper is motivated by the impracticality of individual monitoring of road segments for anomaly detection. Individual monitoring of road segments in large-scale road networks, such as the road network in Manhattan, which consists of more than 8800 road segments is impractical as it leads to a large number of false alarms. Assume the monitoring process of each road segment (using methods such as statistical control charts) is subject to type I error (i.e., false positive) of α . The overall type I error of monitoring all road segments is then calculated as $1 - (1 - \alpha)^n$, where n is the number of roads being monitored. This overall error increases exponentially as n increases. For example, with $\alpha = 0.05$, monitoring a network with only 100 roads results in an overall type I error of 0.994, which results in a very low statistical power. To address this issue, we introduce a temporal self-expressive framework that models the spatiotemporal interdependencies among road segments. This approach detects anomalies by monitoring the changes in the interdependencies among road segments. Road interdependencies are formed due to the spatial adjacency of roads or common travel routes on the network. While our proposed approach can detect anomalies at the road segment level

of granularity, in this study, we group adjacent road segments into small zones. Defining these zones is mainly due to the lack of enough observations in large traffic datasets. That is, traffic data may not always be available for all road segments. These zones are designed to encompass enough data-rich road segments, while also representing the local traffic patterns within their respective regions. Modeling interdependencies among local zones of a traffic network reveals how the traffic flow in each zone relates to the traffic flow in other zones. Furthermore, monitoring the interdependencies over time can help identify unusual changes in traffic patterns and detect anomalous events. In this section, we introduce a temporal self-expressive model to characterize these interdependencies by explaining the traffic in each zone using traffic information from other zones.

Assuming that the traffic network has M zones, we denote the traffic value of zone $i = 1, \dots, M$ at timepoint t by x_{it} and model the data as a multivariate data stream. Using a sliding time window, we arrange the streaming data into a series of equal-size subsequences. In detail, a window of fixed size, L , slides along the data, sample by sample, and creates a series of subsequences such that each subsequence has the same size (i.e., L) and ends one-time step later than the prior subsequence. At a certain timepoint, t , the corresponding subsequence contains traffic data of all zones from timepoint $(t - L + 1)$ to timepoint t , inclusively. Subsequences are numbered sequentially and arranged in two-dimensional arrays $\mathbf{X}_n \in \mathbb{R}^{M \times L}$, where subscript n is the subsequence number ($n = 1, 2, \dots$). For a subsequence $\mathbf{X}_n$, a linear self-expressive model is constructed as follows:

$$x_{it} = \sum_{j \neq i} c_{ij} x_{jt} + e_{it}, \quad \forall \begin{cases} i \in \{1, \dots, M\} \\ t \in T_n \end{cases} \quad (1)$$

where e_{it} represents the model error for zone i at time t , and T_n denotes the set of all timepoints in the given subsequence. At every specific time t , each response variable in the above model (i.e., x_{it}) is also an explanatory variable for all other response variables. Therefore, the model is called self-expressive. Parameters of this model (i.e., c_{ij}) must be estimated in a way that the equality is valid for all $t \in T_n$. This model can be written more compactly by considering $\mathbf{X}_n = [x_{it}] \in \mathbb{R}^{M \times L}$ and defining $\mathbf{C}_n = [c_{ij}] \in \mathbb{R}^{M \times M}$ such that the diagonal elements of $\mathbf{C}_n$ equal zero (i.e., $\text{diag}(\mathbf{C}_n) = 0$). The compact version of the model is written as:

$$\mathbf{X}_n = \mathbf{C}_n \mathbf{X}_n + \mathbf{E}_n, \quad \forall n \in \{1, 2, \dots\} \quad (2)$$

where $\mathbf{E}_n \in \mathbb{R}^{M \times L}$ represents the matrix of residuals. In this model, $\mathbf{C}_n \in \mathbb{R}^{M \times M}$ is a matrix that captures the similarity of zones in terms of their traffic patterns and must be estimated. Since $\mathbf{C}_n$ may become a large matrix when the number of zones is large, its estimation poses significant computational and space complexity challenges. To address these challenges, we impose a low-rank structure on the similarity matrix, $\mathbf{C}_n$, by considering its rank- R decomposition, i.e., $\mathbf{C}_n = \mathbf{U}_n \mathbf{V}_n$, where $\mathbf{U}_n \in \mathbb{R}^{M \times R}$ and $\mathbf{V}_n \in \mathbb{R}^{R \times M}$ ($R \ll M$) are factor matrices. This constraint translates (2) into:

$$\mathbf{X}_n = (\mathbf{U}_n \mathbf{V}_n) \mathbf{X}_n + \mathbf{E}_n, \quad (3)$$

Here, we seek to estimate $\mathbf{U}_n$ and $\mathbf{V}_n$, which requires estimating $2RM$ values instead of M^2 values in (2).

The sliding window moves forward with the passage of time, and the above model is constructed for every consecutive subsequence n ($n = 1, \dots, N$) to characterize the similarity of zones over time. Following the general least squares estimation procedure, the model parameters (i.e., matrix $\mathbf{U}_n$ and matrix $\mathbf{V}_n$) for each subsequence are estimated by solving the following minimization problem under the constraint that $\text{diag}(\mathbf{U}_n \mathbf{V}_n) = 0$:

$$\begin{aligned} \underset{\mathbf{U}, \mathbf{V}}{\text{argmin}} \quad & \frac{1}{2} \|\mathbf{X} - (\mathbf{U} \mathbf{V}) \mathbf{X}\|_F^2 + \alpha_1 \|\mathbf{U}\|_1 + \alpha_2 \|\mathbf{V}\|_1 \\ & + \frac{\beta_1}{2} \|\mathbf{U} - \mathbf{H}\|_F^2 + \frac{\beta_2}{2} \|\mathbf{V} - \mathbf{P}\|_F^2 \\ \text{s.t.} \quad & \text{diag}(\mathbf{U} \mathbf{V}) = 0 \equiv \mathbf{u}^i \mathbf{v}_i = 0 (\forall i \in \{1, \dots, M\}) \end{aligned} \quad (4)$$

In (4), we dropped the subsequence number for ease of notation. In this equation, $\|\cdot\|_1$ is the l_1 norm that induces sparsity in the estimation of the model parameters. Inducing sparsity is favorable as we seek to identify the most informative zones for a given zone. $\mathbf{H}$ and $\mathbf{P}$ are the estimated parameters of the model for the previous subsequence. That is $\mathbf{H} = \mathbf{U}_{n-1}$ and $\mathbf{P} = \mathbf{V}_{n-1}$. We refer to these matrices as history, which are carried over into the current step to control the deviation of the current parameters from their previous values. More specifically, $\|\mathbf{U} - \mathbf{H}\|_F^2$ (or $\|\mathbf{V} - \mathbf{P}\|_F^2$) in the objective function does not allow the estimation of $\mathbf{U}$ (or $\mathbf{V}$) to be arbitrarily different from its history. The motivation for including these penalty terms is that physical traffic does not suddenly jump from one state to another and often has a smooth transition. The constraint $\text{diag}(\mathbf{U} \mathbf{V}) = 0$ eliminates the trivial solution of representing a data point as a linear combination of itself, and α_1 , α_2 , β_1 and β_2 are scalar hyperparameters.

Solving the problem given in (4) is challenging due to the constraint and the non-differentiability of l_1 norms. To solve this problem, we introduce two auxiliary variables which enable the decoupling of the least squared terms (i.e., first, fourth, and fifth terms) from the l_1 norm penalties. Using auxiliary variables $\mathbf{Y}$ and $\mathbf{Z}$, the problem is rewritten as:

$$\begin{aligned} \underset{\mathbf{U}, \mathbf{V}}{\text{argmin}} \quad & \frac{1}{2} \|\mathbf{X} - (\mathbf{U} \mathbf{V}) \mathbf{X}\|_F^2 + \alpha_1 \|\mathbf{Y}\|_1 + \alpha_2 \|\mathbf{Z}\|_1 \\ & + \frac{\beta_1}{2} \|\mathbf{U} - \mathbf{H}\|_F^2 + \frac{\beta_2}{2} \|\mathbf{V} - \mathbf{P}\|_F^2 \\ \text{s.t.} \quad & \mathbf{U} = \mathbf{Y}, \quad \mathbf{V} = \mathbf{Z}, \quad \text{diag}(\mathbf{U} \mathbf{V}) = 0 \end{aligned} \quad (5)$$

Equation (5) is a constrained optimization problem that can be solved by large-scale optimization algorithms such as Alternating Direction Method of Multipliers (ADMM). A full explanation of the ADMM algorithm can be found in [34]. The corresponding augmented Lagrangian function for the problem in (5) is constructed as:

$$\begin{aligned} \mathcal{L} = \quad & \frac{1}{2} \|\mathbf{X} - (\mathbf{U} \mathbf{V}) \mathbf{X}\|_F^2 + \alpha_1 \|\mathbf{Y}\|_1 + \alpha_2 \|\mathbf{Z}\|_1 \\ & + \frac{\beta_1}{2} \|\mathbf{U} - \mathbf{H}\|_F^2 + \frac{\beta_2}{2} \|\mathbf{V} - \mathbf{P}\|_F^2 \\ & + \frac{\rho_1}{2} \|\mathbf{U} - \mathbf{Y}\|_F^2 + \frac{\rho_2}{2} \|\mathbf{V} - \mathbf{Z}\|_F^2 \end{aligned}$$

$$+ \langle \mathbf{A}, \mathbf{U} - \mathbf{Y} \rangle + \langle \mathbf{\Gamma}, \mathbf{V} - \mathbf{Z} \rangle \\ + \sum_{i=1}^M \left(\frac{\rho_3}{2} (\mathbf{u}^i \mathbf{v}_i)^2 + \omega_i \mathbf{u}^i \mathbf{v}_i \right) \quad (6)$$

where $\mathbf{A}$, $\mathbf{\Gamma}$, and ω_i ($i = 1, \dots, M$) are the Lagrangian multipliers (also called dual variables), ρ_1 , ρ_2 and ρ_3 are positive scalars, and $\mathbf{u}^i$ and $\mathbf{v}_i$ are the i^{th} row and column of $\mathbf{U}$ and $\mathbf{V}$, respectively. Lagrangian multipliers $\mathbf{A}$ and $\mathbf{\Gamma}$ have the same dimension as matrices $\mathbf{U}$ and $\mathbf{V}$, respectively.

Using the ADMM algorithm, the problem in (5) is solved by iteratively updating the seven sets of variables $(\mathbf{u}^1, \mathbf{u}^2, \dots, \mathbf{u}^M)$, $\mathbf{Y}$, $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_M)$, $\mathbf{Z}$, $\mathbf{A}$, $\mathbf{\Gamma}$, $(\omega_1, \omega_2, \dots, \omega_M)$. More specifically, at the start of each iteration, rows of matrix $\mathbf{U}$ are updated by fixing all other variables and solving the following problem for each i :

$$\mathbf{u}^i = \underset{\mathbf{u}^i}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{x}^i - \mathbf{u}^i \mathbf{V} \mathbf{X}\|_2^2 + \frac{\beta_1}{2} \|\mathbf{u}^i - \mathbf{h}^i\|_2^2 \\ + \frac{\rho_1}{2} \|\mathbf{u}^i - \mathbf{y}^i\|_2^2 + \langle \lambda^i, \mathbf{u}^i - \mathbf{y}^i \rangle \\ + \frac{\rho_3}{2} (\mathbf{u}^i \mathbf{v}_i)^2 + \omega_i \mathbf{u}^i \mathbf{v}_i \quad (7)$$

where $\mathbf{x}^i$, $\mathbf{h}^i$, $\mathbf{y}^i$ and λ^i are the i^{th} rows of matrices $\mathbf{X}$, $\mathbf{H}$, $\mathbf{Y}$, and $\mathbf{A}$, respectively.

Considering the property of the l_2 norm, $\|\mathbf{a} + \mathbf{b}\|_2^2 = \|\mathbf{a}\|_2^2 + \|\mathbf{b}\|_2^2 + 2 \langle \mathbf{a}, \mathbf{b} \rangle$, the problem in (7) is rewritten as:

$$\mathbf{u}^i = \underset{\mathbf{u}^i}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{x}^i - \mathbf{u}^i \mathbf{V} \mathbf{X}\|_2^2 + \frac{\beta_1}{2} \|\mathbf{u}^i - \mathbf{h}^i\|_2^2 \\ + \frac{\rho_1}{2} \|\lambda^i / \rho_1 + \mathbf{u}^i - \mathbf{y}^i\|_2^2 \\ + \frac{\rho_3}{2} (\mathbf{u}^i \mathbf{v}_i)^2 + \omega_i \mathbf{u}^i \mathbf{v}_i \quad (8)$$

The closed form solution to (8) is derived as follows:

$$\mathbf{u}^i = \left(\mathbf{x}^i \mathbf{X}^T \mathbf{V}^T + \beta_1 \mathbf{h}^i - \lambda^i + \rho_1 \mathbf{y}^i \right. \\ \left. - \omega_i \mathbf{v}_i^T \right) \left(\mathbf{V} \mathbf{X} \mathbf{X}^T \mathbf{V}^T + (\beta_1 + \rho_1) \mathbf{I}_R \right. \\ \left. + \rho_3 \mathbf{v}_i \mathbf{v}_i^T \right)^{-1} \quad (9)$$

where $\mathbf{I}_R$ is the identity matrix of size R .

Next, assuming all other variables are known, auxiliary variable $\mathbf{Y}$ is updated by solving,

$$\mathbf{Y} = \underset{\mathbf{Y}}{\operatorname{argmin}} \alpha_1 \|\mathbf{Y}\|_1 + \frac{\rho_1}{2} \|\mathbf{U} - \mathbf{Y}\|_F^2 + \langle \mathbf{A}, \mathbf{U} - \mathbf{Y} \rangle \\ = \underset{\mathbf{Y}}{\operatorname{argmin}} \alpha_1 \|\mathbf{Y}\|_1 + \frac{\rho_1}{2} \|\mathbf{A} / \rho_1 + \mathbf{U} - \mathbf{Y}\|_F^2. \quad (10)$$

In the second line of (10), the property of the Frobenius norm, $\|\mathbf{A} + \mathbf{B}\|_F^2 = \|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2 + 2 \langle \mathbf{A}, \mathbf{B} \rangle$, is used to simplify the problem.

The closed-form solution to (10) is as [35]:

$$\mathbf{Y} = \operatorname{sgn}(\mathbf{A} / \rho_1 + \mathbf{U}) \circ (|\mathbf{A} / \rho_1 + \mathbf{U}| - \alpha_1 / \rho_1)_+ \quad (11)$$

where $\operatorname{sgn}(a) = a/|a|$ if $a \neq 0$ and $\operatorname{sgn}(a) = 0$ if $a = 0$, and $(a)_+ = \max\{a, 0\}$. These operations, as well as the scalar subtraction, are applied elementwise to matrices. Here,

the symbol $\circ$ denotes the Hadamard product (i.e., elementwise matrix product).

Next, following the same technique that we used to solve for $\mathbf{U}$, the columns of matrix $\mathbf{V}$ are updated by minimizing the following problem for each i :

$$\mathbf{v}_i = \underset{\mathbf{v}_i}{\operatorname{argmin}} \frac{1}{2} \left\| \mathbf{X} - \mathbf{U} \sum_{j \neq i} (\mathbf{v}_j \mathbf{x}^j - \mathbf{U} \mathbf{v}_i \mathbf{x}^i) \right\|_F^2 \\ + \frac{\beta_2}{2} \|\mathbf{v}_i - \mathbf{p}_i\|_2^2 + \frac{\rho_2}{2} \|\mathbf{v}_i - \mathbf{z}_i\|_2^2 \\ + \langle \mathbf{y}_i, \mathbf{v}_i - \mathbf{z}_i \rangle + \frac{\rho_3}{2} (\mathbf{u}^i \mathbf{v}_i)^2 \\ + \omega_i \mathbf{u}^i \mathbf{v}_i \quad (12)$$

where $\mathbf{p}_i$, $\mathbf{z}_i$ and $\mathbf{y}_i$ are the i^{th} columns of matrices $\mathbf{P}$, $\mathbf{Z}$, and $\mathbf{\Gamma}$, respectively.

Solving (12) results in:

$$\mathbf{v}_i = \left(\mathbf{x}^i \mathbf{x}^{iT} \mathbf{U}^T \mathbf{U} + (\beta_2 + \rho_2) \mathbf{I}_R \right. \\ \left. + \rho_3 \mathbf{u}^i \mathbf{u}^{iT} \right)^{-1} \left(\mathbf{U}^T \mathbf{D} \mathbf{x}^{iT} + \beta_2 \mathbf{p}_i - \mathbf{y}_i \right. \\ \left. + \rho_2 \mathbf{z}_i - \omega_i \mathbf{u}^{iT} \right) \quad (13)$$

where $\mathbf{D} = \left(\mathbf{X} - \mathbf{U} \sum_{j \neq i} \mathbf{v}_j \mathbf{x}^j \right)$.

Similar to $\mathbf{Y}$, auxiliary variable $\mathbf{Z}$ is updated by solving,

$$\mathbf{Z} = \underset{\mathbf{Z}}{\operatorname{argmin}} \alpha_2 \|\mathbf{Z}\|_1 + \frac{\rho_2}{2} \|\mathbf{V} - \mathbf{Z}\|_F^2 + \langle \mathbf{\Gamma}, \mathbf{V} - \mathbf{Z} \rangle, \quad (14)$$

which results in:

$$\mathbf{Z} = \operatorname{sgn}(\mathbf{\Gamma} / \rho_2 + \mathbf{V}) \circ (|\mathbf{\Gamma} / \rho_2 + \mathbf{V}| - \alpha_2 / \rho_2)_+ \quad (15)$$

Then, Lagrangian multipliers $\mathbf{A}$ and $\mathbf{\Gamma}$ are updated using the following formulas:

$$\mathbf{A} = \mathbf{A} + \rho_1 (\mathbf{U} - \mathbf{Y}), \quad (16)$$

$$\mathbf{\Gamma} = \mathbf{\Gamma} + \rho_2 (\mathbf{V} - \mathbf{Z}), \quad (17)$$

Finally, Lagrangian multipliers ω_i ($i = 1, \dots, M$) are updated for each i as:

$$\omega_i = \omega_i + \rho_3 (\mathbf{u}^i \mathbf{v}_i) \quad (18)$$

It should be noted that in the above update rules, the most recent estimates of the variables are used. Iterating the above procedure decreases the loss function in (5) together with the residuals of the constraints and leads to a solution for $\mathbf{U}$ and $\mathbf{V}$ such that either they cannot be improved any further by the updating rules or a maximum number of iterations is reached.

Having matrices $\mathbf{U}$ and $\mathbf{V}$ estimated, the similarity matrix, $\mathbf{C}$, is estimated as:

$$\hat{\mathbf{C}} = \mathbf{U} \mathbf{V} \quad (19)$$

The pseudocode of the ADMM algorithm for the proposed self-expressive model is outlined in Algorithm 1.

In implementing Algorithm 1, we use two stopping criteria and continue with the updates until one is met. Firstly, we check whether both $\mathbf{U}$ and $\mathbf{V}$ matrices have converged.

Algorithm 1 ADMM Algorithm for Solving (5)

input matrix X , history matrices H, P , parameters $\alpha_1, \alpha_2, \beta_1, \beta_2, \rho_1, \rho_2, \rho_3$ and rank R
initialize $U = Y = 0, V = Z, \mathbf{A} = 0, \mathbf{\Gamma} = 0$, and $\omega_i = 0 (i \in \{1, \dots, M\})$
repeat
 for $i \in \{1, \dots, M\}$ **do**
 update u^i by (9)
 end for
 update Y by (11)
 for $i \in \{1, \dots, M\}$ **do**
 update v_j by (13)
 end for
 update Z by (15)
 update $\mathbf{A}$ by (16)
 update $\mathbf{\Gamma}$ by (17)
 for $i \in \{1, \dots, M\}$ **do**
 update ω_i by (18)
 end for
until convergence criterion or maximum iterations is reached
return U, V

The algorithm terminates if both $\|U^{(k)} - U^{(k-1)}\|_F < \varepsilon$ and $\|V^{(k)} - V^{(k-1)}\|_F < \varepsilon$ are satisfied, where k represents the iteration number, and ε is the tolerance. Secondly, we set a maximum number of iterations, which can be adjusted based on the user's computational resources.

B. Computational Complexity

In this section, we analyze the computational complexity of the proposed temporal self-expressive model. The computational complexity of calculating u^i using Equation (9) is $O(LMR + R^2L + \frac{1}{3}R^3 + R^2)$, which reduces to $O(LMR + R^2L + \frac{1}{3}R^3)$. Since $R \ll M$, the complexity further reduces to $O(LMR + \frac{1}{3}R^3)$. Note that Equation (9) requires the inversion of an $R \times R$ matrix. The standard Gauss-Jordan elimination requires $O(R^3)$ steps. However, this can be reduced to $O(\frac{1}{3}R^3)$ steps with Cholesky decomposition for $R > 1$. This procedure is performed for every $u^i (i = 1, \dots, M)$. Thus, the update of U requires $O(LM^2R + \frac{1}{3}R^3M)$ steps. The computational complexity of updating Y using Equation (11) is $O(MR)$. The computational complexity of updating v_i using Equation (13) is $O(L + R^2M + \frac{1}{3}R^3 + LMR + R^2)$, which reduces to $O(R^2M + LMR)$. Therefore, the update of matrix V requires $O(R^2M^2 + LM^2R)$ steps. The computational complexity of updating Z using Equation (15) is $O(MR)$. Finally, the updates of $\mathbf{A}, \mathbf{\Gamma}$, and $\omega_i (i = 1, \dots, M)$ require $O(MR), O(MR)$, and $O(MR)$ steps, respectively. Therefore, the complexity of a single ADMM iteration is calculated as $O(LM^2R + \frac{1}{3}R^3M + MR + R^2M^2 + LM^2R + MR + MR + MR + MR)$, which reduces to $O(LM^2R + R^2M^2)$. Assuming K iterations of the ADMM algorithm is performed to estimate all variables, the overall complexity of the proposed method is $O(KR^2M^2 + KLM^2R)$.

C. Statistical Monitoring for Anomaly Detection

The temporal self-expressive model proposed in the previous section sets the stage for establishing a monitoring framework to detect network-wide traffic anomalies and identify the local zones most affected by the anomalous event. As described in Section III-A, this model expresses the traffic of each zone using traffic information from other zones. During normal conditions, the model is developed to capture regular interdependencies among traffic in different zones. Therefore, when applied to regular traffic, it can reliably explain the traffic condition of each zone. However, when an extreme event impacts the network, traffic interdependencies between local zones change significantly, and the developed model cannot account for unusual patterns of traffic. This leads to large and dispersed errors in the model. In view of this fact, we can detect network-wide anomalies by monitoring a statistic that measures the goodness of fit of the self-expressive model over time.

At each timepoint, t , we estimate the parameters of the self-expressive model (i.e., matrix $\hat{C}$) using a subsequence that spans from $(t - L + 1)$ to t . With these estimated parameters, we can compute the error vector of the model as $e_t = \hat{C}x_t - x_t$, where $x_t \in \mathbb{R}^M$ contains traffic values of all zones at timepoint t . If the model captures most of the variation within the traffic data, the resulting errors will be small with limited variance compared to the original variance of the data. Therefore, we define,

$$G_t = \frac{\text{var}(e_t)}{\text{var}(x_t)}, \quad (20)$$

as a statistic that measures the quality of fit of the self-expressive model. Larger values of G indicate that the model could not capture the patterns within the traffic data, meaning that different zones cannot explain each other. This situation often appears when an anomalous event occurs in the network.

Our monitoring scheme is based on statistical process control (SPC) charts. Control charts are powerful monitoring tools used to detect statistically significant changes or deviations in a process from its expected pattern. A typical control chart plots a certain statistic against time and uses upper and lower control limits (UCL and LCL) to identify abnormal variations. These control limits are determined based on the typical variation of the monitoring statistic obtained from historical data. An out-of-control signal is raised when the statistic goes beyond the control limits, indicating the presence of a source of anomaly. Montgomery [36] provides a detailed description of different types of control charts.

In this study, we create an empirical control chart to monitor the goodness-of-fit for the self-expressive model. The chart is developed in two phases: Phase 1 and Phase 2. In Phase 1, we utilize historical data to determine the underlying probability distribution of the monitoring statistic, G . Using this distribution, we then set the control limits of the chart based on a desired significance level. This significance level serves as an upper bound for the probability of observing a G value that exceeds the control limits. By adjusting the significance level, the method can detect events of different severity; a lower level can focus on the most severe events, while a

higher level can capture less severe events at the cost of more frequent false alarms. In Phase 2, we plot the monitoring statistic, G , over time. If traffic condition deviates from its typical pattern, the statistic will fall outside of the control limits, indicating the occurrence of an unusual event in the network. By implementing this approach, we can reliably monitor network performance and detect large-scale anomalies in a timely manner.

D. Identification of the Anomalous Zones

After detecting anomalous events at the network level, our second goal is to perform zone-level diagnostics to identify the zones that exhibit the most anomalous traffic patterns. Notably, we examine the error vectors of the self-expressive model, i.e., $\mathbf{e}_t$, which allows us to gain insights into the traffic behavior of each zone at every point in time. When the error value for a particular zone i at timepoint t , i.e., e_{it} , becomes large, it indicates that the zone is experiencing an abnormal traffic pattern that the model cannot explain. In other words, the traffic pattern of the zone cannot be explained by the other zones in the network. A naïve approach to identify the zones with anomalous behavior is to monitor the model error of each zone individually. However, in this approach, performance of the monitoring system drops significantly due to high false alarm rate. To address this problem, we propose to calculate anomaly scores at the zone level based on the error vectors of the self-expressive model.

Considering that the model errors of all M zones at timepoint t are represented by a vector $\mathbf{e}_t \in \mathbb{R}^M$, we obtain a sparse representation of this vector by imposing a lasso penalty as follows:

$$\underset{\mathbf{s}_t}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{e}_t - \mathbf{s}_t\|^2 + \theta \|\mathbf{s}_t\|_1 \quad (21)$$

where $\mathbf{s}_t$ denotes the sparse representation of vector $\mathbf{e}_t$.

Equation (21) can be solved by applying a soft-thresholding operator to the vector $\mathbf{e}_t$ [37] as:

$$\mathbf{s}_t = \operatorname{sgn}(\mathbf{e}_t) \circ (|\mathbf{e}_t| - \theta)_+ \quad (22)$$

where $\operatorname{sgn}(\cdot)$ is the sign function, and $(a)_+$ represents the positive part of a real number a which is defined as $(a)_+ = \max\{a, 0\}$. The soft-thresholding operator is applied elementwise to the vector $\mathbf{e}_t$.

The absolute values of the sparse representation $\mathbf{s}_t$ are used as anomaly scores. Specifically, the absolute value of the i th element of vector $\mathbf{s}_t$ represents the anomaly score for zone i at timepoint t . The larger the score, the more anomalous the behavior in zone i is at timepoint t .

We select the threshold θ such that the anomaly score for the majority of the zones (at least 80% of the zones) becomes zero during non-event periods. This scoring approach enables us to both identify the zones with anomalous traffic behavior, i.e., $\{i | s_{it} > 0\}$, and find the zones with the highest potential of causing a change in the network by ranking the non-zero scores. When an extreme event occurs, the proposed approach allows us to identify and monitor zones with high anomaly scores and take actions accordingly to improve overall traffic flow.

IV. APPLICATION TO HURRICANE SANDY IN NEW YORK CITY

In this section, we assess the effectiveness of the proposed self-expressive network monitoring method by applying it to the road network of Manhattan in New York City. Specifically, we evaluate its ability to detect traffic anomalies caused by Hurricane Sandy in 2012.

A. Dataset Description

In this study, we analyze traffic on the New York City road network using a publicly available dataset derived from taxi trajectories [38]. Many previous studies, including [8], [39], and [40], examined the reliability of trajectory data as a proxy that represents traffic patterns in road networks and used it in various applications. The New York City traffic dataset was estimated from records of nearly 700 million taxi trips in New York City over the years 2010-2013. This dataset contains hourly average travel times on individual road segments of the NYC road network from January 2010 to December 2013. It also includes information about the coordinates and length of each road segment. Previous studies showed that many events occurred within these four years [8], [23]. Among these events, this study pays attention to Hurricane Sandy, which formed on October 22, 2012, and hit New York City on October 29, 2012.

For the purpose of this study, the traffic data for a period of three weeks, from Monday, October 15, 2012, to Monday, November 5, 2012, in the borough of Manhattan is selected. The road network in Manhattan consists of 8839 road segments, represented in the dataset by 3910 nodes and 8839 links.

B. Data Preprocessing

Since the traffic dataset is incomplete, in the first step, we group adjacent road segments by their geographical locations. Specifically, the Manhattan area is divided into forty-six small equally-sized zones (Fig. 2). A road segment is assigned to a zone if its midpoint falls within the boundaries of that zone. From a theoretical perspective, we can continue reducing the size of traffic zones to the point where each zone encompasses only a single road segment. This approach allows us to create a model that can detect anomalous traffic at the road segment level. However, the quality of such models can be severely affected by the lack of observations in many road segments, particularly during off-peak hours. Dividing the road network area into equally-sized zones is an appropriate approach for the Manhattan area, considering the homogeneity of the traffic network. However, as discussed in the conclusion section, further investigation on determining the shape and size of the zones in other urban areas can be a topic for future studies.

To identify unusual traffic patterns and detect anomalies, each zone's hourly average pace of traffic (i.e., travel time per unit of distance) is used in the proposed self-expressive model. Using pace as the key traffic metric allows for accommodating the varied length of road segments within zones [8]. On a single road segment, the average pace of traffic is equal

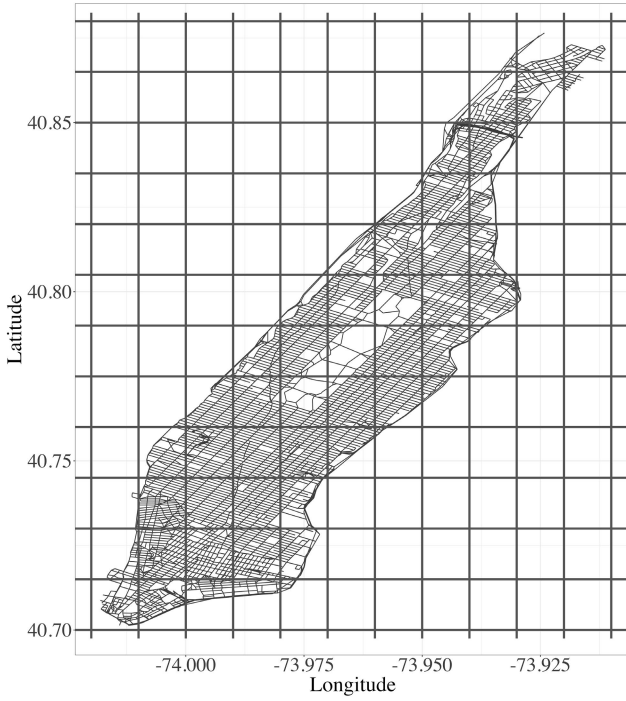


Fig. 2. Partitioning of the road network of Manhattan into forty-six zones.

to the average travel time divided by the road segment's length, which is simply the inverse of the average speed [41]. In this paper, the average pace of a zone is defined as the length-weighted average of all its road segments' paces. More precisely, the average pace of zone i at time t (i.e., x_{it}) is calculated as:

$$x_{it} = \frac{\sum_{r \in S_i} l(r)p(r, t)}{\sum_{r \in S_i} l(r)} \quad (23)$$

where S_i denotes the set of all road segments in zone i , $p(r, t)$ is the average pace of traffic on road segment r at time t , and $l(r)$ is the length of road segment r .

Hurricane Sandy hit New York City on October 29, 2012. Therefore, we use traffic data recorded from October 15, 2012, to November 5, 2012, for analysis. Fig. 3 shows the average pace of traffic in a sample zone over the selected period. The figure reveals the expected periodic pattern of traffic data. For example, during rush hours, when the speed of traffic is lower, the pace is higher, whereas at night (when typically, the speed is higher), the pace values are lower. Moreover, rush hour is less extreme on weekends. During regular days, there may be a small variance around these patterns, but, as it is illustrated in the figure, the pace varies significantly when an unusual event occurs.

The hourly average pace of all zones can be viewed as a multivariate data stream. A representation of this data stream is a matrix of size $M \times T$, where $M = 46$ is the number of zones in the network and $T = 504$ is the number of timepoints in the data stream (i.e., $24 \times 7 \times 3 = 504$ hours). This matrix contains the hourly average pace values of all the forty-six zones from October 15, 2012, to November 5, 2012.

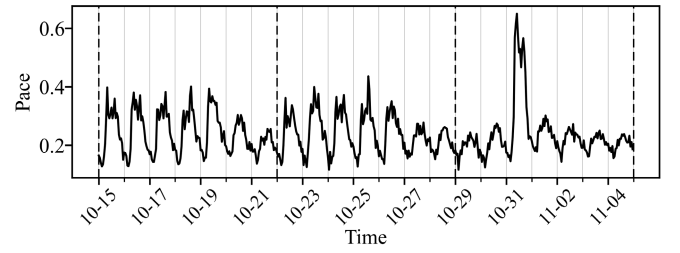


Fig. 3. Hourly average pace of traffic in a sample zone during the selected three-week period (October 15, 2012, to November 5, 2012).

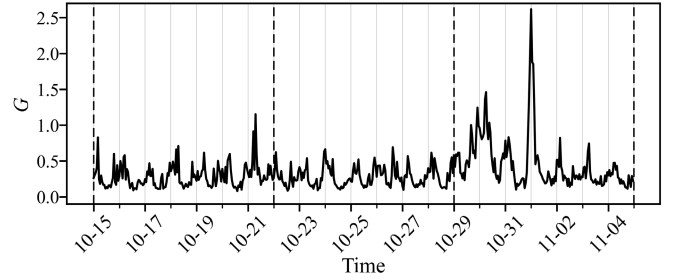


Fig. 4. Demonstration of the changes in monitoring statistic G over the selected three-week period (October 15, 2012, to November 5, 2012).

C. Selection of Hyperparameters

Algorithm 1 requires the selection of parameters α_1 , α_2 , β_1 , β_2 , ρ_1 , ρ_2 , ρ_3 and rank R . In our experiment, we choose a low rank R of 3. For ease of computation, α_1 and α_2 can be set to the same value, and we set $\alpha_1 = \alpha_2 = 1/(2\sqrt{\max\{M, R\}}) = 0.0737$. We follow the same idea for β_1 and β_2 and set them to the same value. Specifically, we determine β_1 and β_2 as the inverse of the average value of $\|x_t - x_{t-1}\|_2$ for every $1 \leq t \leq 168$ (data prior to October 22, 2012).

In addition, we set the starting values of ρ_1 , ρ_2 and ρ_3 to 10^{-1} , which are increased geometrically over the iterations by a fixed factor $f = 2$ up to 10^4 . Moreover, we set the tolerance ε to 10^{-5} and the maximum number of iterations to 10^3 . Lastly, we choose the window size $L = 8$ for casting the data stream into subsequences. Generally, traffic patterns in an urban area show three different regimes almost every 8 hours (i.e., morning, afternoon, and evening). Therefore, in this study, we selected a window size of 8 to ensure that the model effectively captures these traffic patterns.

D. Network-Level Anomaly Detection

In order to detect traffic anomalies at network level, we first obtain the parameters of the self-expressive model at every timepoint over the selected three-week period. In more detail, the self-expressive parameters (i.e., matrix $\hat{C}$) for each hour, t , are estimated through Algorithm 1 using a subsequence of the data that spans from hour $(t - L + 1)$ to hour t , where $L = 8$. Next, using the estimated parameters, we obtain the error vector (i.e., $e_t = \hat{C}x_t - x_t$) of the model at every timepoint t . We then use these error vectors to quantify the quality of fit of the model (i.e., G_t in (20)) over the selected analysis period.

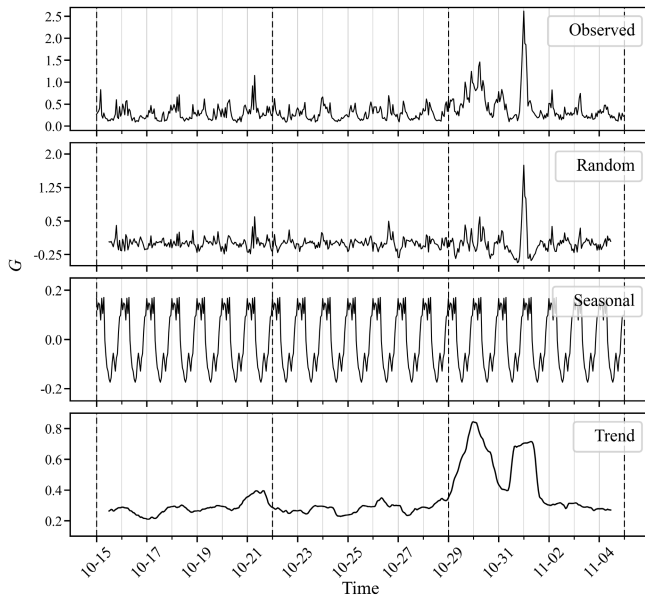


Fig. 5. Decomposition process to remove seasonality patterns from statistic G .

Fig. 4 shows the resulting time series of monitoring statistic G over the three-week analysis period. Unusual changes in G during the week of the hurricane are visually apparent in this figure, suggesting that the traffic has deviated from its normal pattern because of an unusual event.

As Fig. 4 shows, statistic G exhibits a daily seasonal pattern. Notably, during post-midnight hours, when road traffic is reduced and interdependencies among traffic is weaker, G values are generally higher. To account for this daily pattern, we apply a moving average technique [42] with a window size of 24 hours to extract the underlying trends in the data. Fig. 5 illustrates the process of removing the seasonality and extracting the trends. The extracted trends are further monitored using a statistical control chart to accurately identify atypical patterns and detect traffic anomalies at the network level.

The statistical control chart is created based on the underlying distribution of our monitoring statistic. Specifically, we analyze historical data from a period characterized by normal traffic conditions to empirically determine the distribution of the monitoring statistic. Based on this empirical distribution, we determine the upper and lower control limits (UCL and LCL) for the chart to achieve a desired significance level.

In this study, we employ the Python library *distfit* [43] to identify the distribution that best fits the monitoring statistic during normal conditions (prior to October 22, 2012). A Weibull distribution, with shape = 0.881, location = 0.285 and scale = 0.032 parameters, is identified as the best fit (as shown in Fig. 6). Using this distribution, we set the upper and lower control limits at a 0.02 significance level to detect abnormal traffic patterns. The calculated UCL and LCL are 0.404 and 0.165, respectively (depicted as red dashed lines in Fig. 6). It is worth noting that a lower significance level would result in wider control limits and reduce the number of

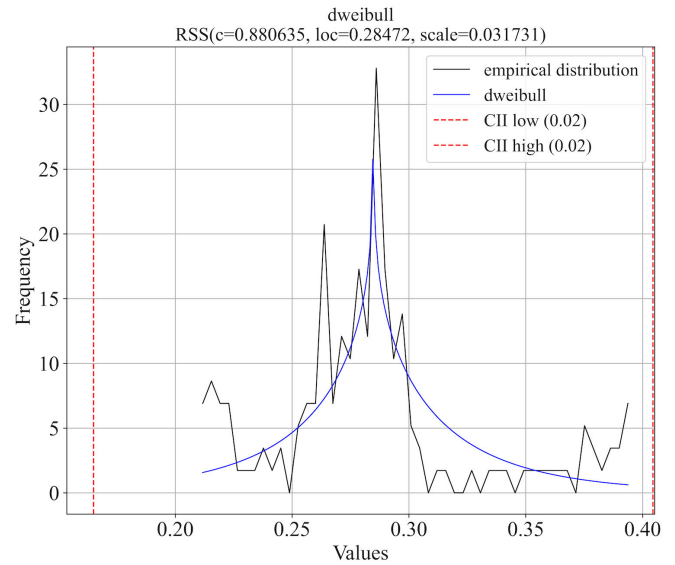


Fig. 6. Best-fitted distribution to the data in Phase 1 of the empirical control chart.

false alarms at the cost of overlooking small shifts. A higher significance level would yield tighter control limits and enable detection of smaller anomalous shifts at the cost of more frequent false alarms.

Fig. 7 shows phases 1 and 2 of the empirical control chart. As the figure illustrates, traffic anomalies are detected on October 29, 2012, when Hurricane Sandy struck the NYC area. The anomalies start after midnight on October 29, 2012, and last until midday on November 1, 2012. The figure demonstrates that the proposed framework can accurately capture the abnormal traffic patterns at the network level.

E. Zone-Level Anomaly Detection

To analyze how Hurricane Sandy affected traffic behavior across different zones, we calculate an anomaly score for each zone at every hour over the analysis period. As explained in Section III-C, these scores are calculated based on the error vectors of the self-expressive model (i.e., e_t). In this study, we first use a moving average technique [42] with a window size of 24 hours to extract the hourly vectors of the trends in the model errors of individual zones. The underlying reason is to remove seasonal patterns from the error sequences. We then apply a soft-thresholding operator to these vectors to obtain the anomaly scores of every zone over time. For this study, we set the threshold parameter θ to 0.01, resulting in 82% of the scores being zero during the week before the hurricane (non-event period). Our goal is to identify the zones whose traffic is affected by the hurricane the most.

Fig. 8(a) shows the hourly anomaly scores for all forty-six zones from October 22, 2012, to November 5, 2012. The anomaly scores reflect the degree of deviation from expected behavior, with a higher score indicating more anomalous traffic behavior in a specific zone at a given time. The network disturbance caused by the hurricane is evident in this figure. As shown in the figure, certain zones exhibit a higher level of abnormality, which is expected during an extreme event.

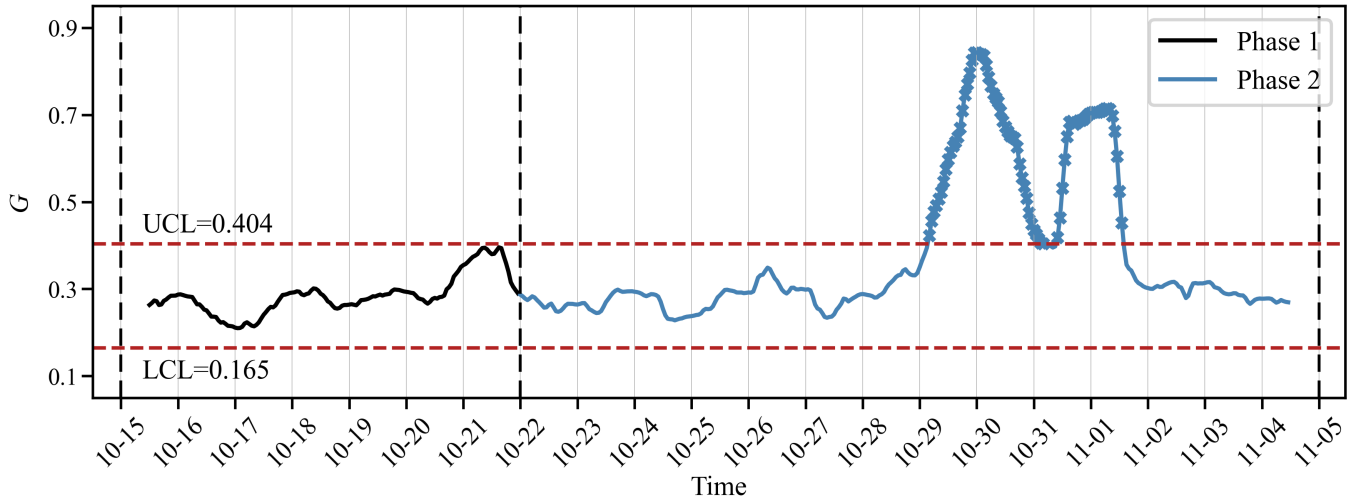


Fig. 7. Empirical control chart of statistic G detecting anomalies on October 29, 2012.

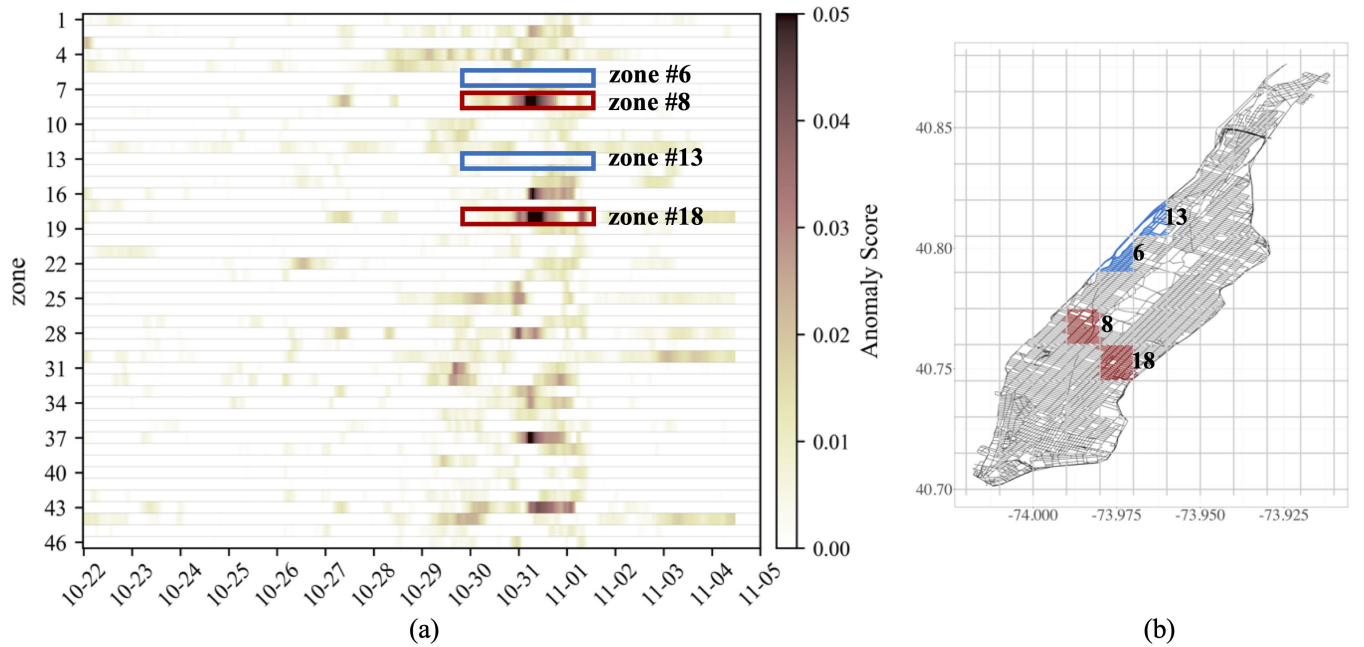


Fig. 8. (a) Anomaly scores for every zone from October 22, 2012, to November 5, 2012. During the hurricane, zones #8 and #18 have the highest anomaly scores, while zones #6 and #13 receive the lowest anomaly scores. (b) locations of the identified zones on the Manhattan network.

Depending on their location, population and infrastructure, some zones may experience more anomalous traffic patterns than others.

To demonstrate the advantages and effectiveness of the proposed method, we use a face validation approach, in which the traffic condition (pace) of a road is visually examined for various zones with different anomaly scores (computed by our method). The results are demonstrated in Figure 8 and Figure 9. The outcomes confirm that the zones with a higher anomaly score show stronger shifts in their pace time-series data, whereas the ones with a lower score show negligible changes in the pace. More specifically, Fig. 8(a) demonstrates zones #8 and #18 received the highest anomaly scores on October 31. This observation is validated by their pace plots

in Fig. 9(a) and 9(b), which show severe congestion in these zones during that time. In contrast, zones #6 and #13 have low anomaly scores during the hurricane, indicating that they were less affected by the network disturbance. Their pace plots in Fig. 9(c) and 9(d) confirm that they did not experience significant disruption during the hurricane. Fig. 8(b) illustrates the locations of these zones on the Manhattan network. As shown on the map, the two most anomalous zones (i.e., zones #8 and #18) are located in midtown, while zones with the lower scores are both in the upper west side of Manhattan.

F. Sensitivity Analysis

In this section, we examine the sensitivity of our proposed method to the hyperparameters α_1 , α_2 , β_1 , β_2 , rank R , and

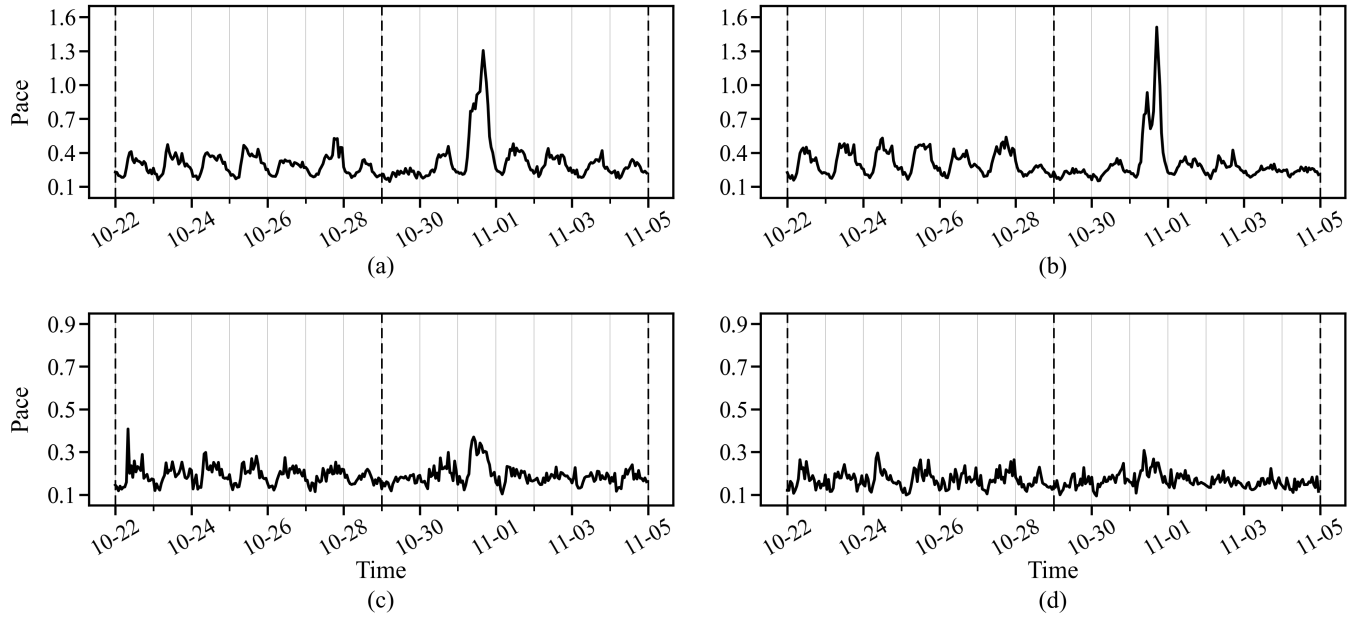


Fig. 9. Actual pace plots of zones with high ((a) and (b)) and low ((c) and (d)) anomaly scores. (a) zone #8. (b) zone #18. (c) zone #6. (d) zone #13.

TABLE II
SENSITIVITY TO THE RANK R AND WINDOW SIZE L

(R, L)	(3, 8)	(5, 8)	(7, 8)	(4, 12)	(6, 12)	(8, 12)	(8, 24)	(12, 24)	(14, 24)
τ (hour)	4	4	4	5	4	5	5	4	5

window size L . First, we consider the time between the first detected anomaly and the beginning of October 29 (the day that Hurricane Sandy struck NYC area) as an evaluation metric for the sensitivity analysis. We denote this metric by τ . Next, we assess the performance of the proposed method in terms of τ , for different combinations of the hyperparameters.

In the proposed method, α_1 and α_2 control the sparsity of self-expressive parameters, while β_1 and β_2 regulate the deviations of the model parameters at each timepoint from their preceding values. As detailed in Section IV-C, following the theoretical results in [44], we set α_1 and α_2 to $1/(2\sqrt{\max\{M, R\}})$, where M denotes the number of zones, and R is the rank of the similarity matrix, $\mathbf{C}$. Additionally, β_1 and β_2 were both set to the inverse of the average norm of the difference between pace vectors at consecutive timepoints ($\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2$) for data prior to October 22, 2012 (i.e., $1 \leq t \leq 168$). This value captures the average variations in traffic pace under normal conditions and provides a reasonable choice for balancing the deviations of the model parameters from their previous values. Specifically, in our study, we set $\alpha_1 = \alpha_2 = \alpha = 0.0737$ and $\beta_1 = \beta_2 = \beta = 3.50$.

To explore the sensitivity to the hyperparameters α and β , we vary α between 0.0537 and 0.0937, and β between 2.50 and 4.5. The remaining hyperparameters (i.e., rank R , and window size L) are set as described in Section IV-C. Fig. 10 shows τ for different combination of α and β . The results indicate that lower values of β , along with higher values of α , result

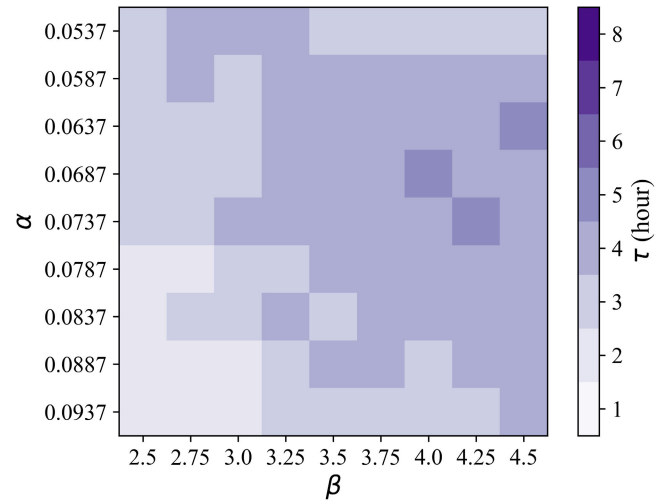


Fig. 10. Sensitivity analysis for the hyperparameters α and β .

in a slightly quicker anomaly detection. However, overall, the proposed anomaly detection method is not highly sensitive to these two hyperparameters, meaning that small changes in their values do not significantly impact the anomaly detection outcomes.

Next, we examine the sensitivity of our method to rank R and window size L . Table II reports the values of τ for different combinations of R and L , indicating that the

proposed anomaly detection method is not considerably sensitive to these two hyperparameters. While the sensitivity to L and R is not significant, one should select the window size and the rank sufficiently large to encompass a large enough sample of traffic patterns and effectively capture the complexity of the traffic patterns over the selected window.

To examine the proposed method more rigorously, we applied the method to another extreme event (a blizzard that hit New York City on December 26, 2010) and presented the outcomes in the supplementary document. Like the results of Hurricane Sandy, the anomaly scores computed using the proposed method agree with the visual examinations of the traffic patterns measured by pace [45].

G. Discussion

Our results show that the proposed framework can accurately detect anomalous traffic patterns at both network and zone levels. In comparison to existing anomaly detection methods that analyze traffic for each road segment individually, which may lead to a significant number of false alarms in large-scale networks, our proposed method efficiently monitors the entire road network and identifies the most anomalous zones by imposing a sparsity-inducing penalty on the errors of the self-expressive model. By adopting the proposed approach, decision-makers can determine which regions of the network are most affected by the extreme event and prioritize their mitigation efforts accordingly.

The current implementation of our proposed method creates zones that consist of a sufficient number of data-rich road segments. This approach allows us to mitigate the impact of incomplete data. However, in cases where enough data is not available for certain zones, we can employ data imputation techniques to handle incomplete data. There is a body of research that has developed methods for imputing traffic data in large-scale networks (e.g., [46], [47], [48]). These methods can be integrated into our current work to address potential problems of incomplete data.

The proposed method offers flexibility in defining traffic zones, without specific constraints on their shape or the type and number of roads they may contain. In this study, given the relatively homogeneous structure of Manhattan's road network, we partitioned the network into grids of comparable size that resulted in zones with relatively similar sets of road segments. It is worth noting that the proposed self-expressive model can capture the collective explanatory power of each zone to model traffic patterns in other zones. Therefore, if the road segments in one zone significantly differ from those in another zone in a manner that leads to a lower level of explanatory power, the self-expressive model can capture these distinctions and provide reliable outcomes. This adaptable approach enables our method to account for the varying road types and traffic levels throughout the network. In addition, our method is specifically designed to identify events that create statistically significant changes in traffic patterns. This enables us to distinguish shifts in the traffic data from background noise.

Furthermore, our proposed method has the potential to be integrated into an online traffic monitoring system. The

method can continuously receive and analyze traffic data streams to detect abnormalities in traffic patterns. In addition, the method's ability to constantly adapt and learn from evolving traffic conditions makes it a reliable tool for traffic management during extreme events.

The results from the computational complexity analysis in Section III-B indicate that the computational cost of the proposed method increases quadratically as the number of zones (denoted by M) in a network increases. Therefore, the proposed method is scalable to real-world road networks. Furthermore, a hierarchical approach where the impacted area is first divided into small number of zones and then a zone that shows abnormal traffic patterns is divided into smaller zones, can be used to improve the overall scalability of the method.

Finally, data privacy in traffic monitoring may cause concerns. To address this issue, the proposed method is designed to rely exclusively on the aggregation and analysis of vehicle speeds on road segments, without utilizing any form of personally identifiable information (PII) derived from the vehicles.

V. CONCLUSION

Timely and reliable detection of traffic anomalies in road networks is a critical step toward developing resilient transportation systems. The existing data-driven methods for road traffic monitoring during extreme events are limited to detecting anomalies at the network level and are not capable of identifying anomalous traffic at the zone level. This study aims to address this limitation by proposing a temporal self-expressive network monitoring method that can detect anomalies caused by extreme traffic events and determine what parts (zones) of the network are most affected by the anomalous event. In this method, first, we divide the road network into small zones. Next, we develop a temporal self-expressive model to characterize the dynamic traffic interdependencies between these zones. Then, we employ a statistical monitoring tool to detect large-scale traffic anomalies caused by extreme events through monitoring the goodness of fit of the self-expressive model. Finally, we calculate an anomaly score for every zone over time to determine which zones are considerably affected by the extreme event.

We applied our proposed method to the road network of Manhattan in New York City to evaluate its performance in detecting traffic anomalies caused by Hurricane Sandy in 2012 and a major blizzard in 2010. The results confirmed that the proposed self-expressive network monitoring method could effectively detect anomalous traffic resulting from the events at both network and zone levels.

The primary contribution of this study to the existing core body of knowledge is to create a network monitoring method that can detect traffic anomalies during extreme events at both network and zone levels. Further investigations to determine the shape and size of zones, especially in non-homogenous urban areas, can be a potential topic for future studies. Investigations to address the problem of incomplete data and develop practical anomaly detection solutions at a road-segment level of granularity can be another basis for future research.

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] R. Blumenthal. (Sep. 2005). *Miles of Traffic as Texans Heed Order to Leave*. New York Times. [Online]. Available: <https://www.nytimes.com/2005/09/23/us/nationalspecial/miles-of-traffic-as-texans-heed-order-to-leave.html>
- [2] M. Levin. (Aug. 2017). *How Hurricane Rita Anxiety Led to the Worst Gridlock in Houston History*. [Online]. Available: <https://www.chron.com/news/houston-texas/houston/article/Hurricane-Rita-anxiety-leads-to-hellish-fatal-6521994.php>
- [3] S. W. Khan et al., "Anomaly detection in traffic surveillance videos using deep learning," *Sensors*, vol. 22, no. 17, p. 6563, Aug. 2022.
- [4] J. Ren, Y. Chen, L. Xin, J. Shi, B. Li, and Y. Liu, "Detecting and positioning of traffic incidents via video-based analysis of traffic states in a road segment," *IET Intell. Transp. Syst.*, vol. 10, no. 6, pp. 428–437, Aug. 2016.
- [5] E. P. Ijjina, D. Chand, S. Gupta, and K. Goutham, "Computer vision-based accident detection in traffic surveillance," in *Proc. 10th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Kanpur, India, Jul. 2019, pp. 1–6.
- [6] F. Harrou, A. Zeroual, and Y. Sun, "Traffic congestion monitoring using an improved kNN strategy," *Measurement*, vol. 156, May 2020, Art. no. 107534.
- [7] K. Kalair and C. Connaughton, "Anomaly detection and classification in traffic flow data from fluctuations in the flow–density relationship," *Transp. Res. C, Emerg. Technol.*, vol. 127, Jun. 2021, Art. no. 103178.
- [8] 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.
- [9] W. Kuang, S. An, and H. Jiang, "Detecting traffic anomalies in urban areas using taxi GPS data," *Math. Problems Eng.*, vol. 2015, pp. 1–13, Jan. 2015.
- [10] X. Han, T. Grubenmann, R. Cheng, S. C. Wong, X. Li, and W. Sun, "Traffic incident detection: A trajectory-based approach," in *Proc. IEEE 36th Int. Conf. Data Eng. (ICDE)*, Apr. 2020, pp. 1866–1869.
- [11] S. Xu, S. Li, W. Huang, and R. Wen, "Detecting spatiotemporal traffic events using geosocial media data," *Comput., Environ. Urban Syst.*, vol. 94, Jun. 2022, Art. no. 101797.
- [12] K. C. Roy, M. Cebrian, and S. Hasan, "Quantifying human mobility resilience to extreme events using geo-located social media data," *EPJ Data Sci.*, vol. 8, no. 1, pp. 1–15, May 2019.
- [13] S. Ahmad, A. Ali, H. U. Ahmed, Y. Huang, and P. Lu, "Evaluating traffic operation conditions during wildfire evacuation using connected vehicles data," *Fire*, vol. 6, no. 5, p. 184, May 2023.
- [14] V. C. Kummetha, M. Kamrani, S. Concas, A. Kourtellis, and O. Dokur, "Proactive congestion management via data-driven methods and connected vehicle-based microsimulation," *J. Intell. Transp. Syst.*, vol. 2022, pp. 1–17, Nov. 2022.
- [15] R. Bauza and J. Gozalvez, "Traffic congestion detection in large-scale scenarios using vehicle-to-vehicle communications," *J. Netw. Comput. Appl.*, vol. 36, no. 5, pp. 1295–1307, Sep. 2013.
- [16] 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.
- [17] 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, 2019, pp. 625–633.
- [18] 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.
- [19] 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.
- [20] J. Xiao, "SVM and KNN ensemble learning for traffic incident detection," *Phys. A, Stat. Mech. Appl.*, vol. 517, pp. 29–35, Mar. 2019.
- [21] 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.
- [22] 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.
- [23] 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.
- [24] N. Dogru and A. Subasi, "Traffic accident detection using random forest classifier," in *Proc. 15th Learn. Technol. Conf.*, Jeddah, Saudi Arabia, Feb. 2018, pp. 40–45.
- [25] L. Li, Y. Lin, B. Du, F. Yang, and B. Ran, "Real-time traffic incident detection based on a hybrid deep learning model," *Transp. A, Transp. Sci.*, vol. 18, no. 1, pp. 78–98, 2022.
- [26] 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)*, Maui, HI, USA, Nov. 2018, pp. 1011–1016.
- [27] 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.
- [28] M. Ilbeigi, "Statistical process control for analyzing resilience of transportation networks," *Int. J. Disaster Risk Reduction*, vol. 33, pp. 155–161, Feb. 2019.
- [29] Y. Hu, A. Qu, and D. Work, "Detecting extreme traffic events via a context augmented graph autoencoder," *ACM Trans. Intell. Syst. Technol.*, vol. 13, no. 6, pp. 1–23, Sep. 2022.
- [30] S. Yang, K. Kalpakakis, and A. Biem, "Detecting road traffic events by coupling multiple timeseries with a nonparametric Bayesian method," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 1936–1946, Oct. 2014.
- [31] Q. Li, H. Tan, Z. Jiang, Y. Wu, and L. Ye, "Nonrecurrent traffic congestion detection with a coupled scalable Bayesian robust tensor factorization model," *Neurocomputing*, vol. 430, pp. 138–149, Mar. 2021.
- [32] 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, Dec. 2020.
- [33] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev. Soc. Ind. Appl. Math.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.
- [34] S. Boyd et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
- [35] E. T. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for ℓ_1 -minimization: Methodology and convergence," *SIAM J. Optim.*, vol. 19, no. 3, pp. 1107–1130, 2008.
- [36] D. C. Montgomery, *Introduction to Statistical Quality Control*. Hoboken, NJ, USA: Wiley, 2009.
- [37] T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical Learning With Sparsity: The Lasso and Generalizations*. Boca Raton, FL, USA: CRC Press, 2015.
- [38] B. Donovan and D. B. Work. (Jun. 2015). *Open Data: 2010–2013 New York City Traffic Estimates*. Univ. Illinois. [Online]. Available: <https://uofi.app.box.com/v/NYC-traffic-estimates>
- [39] J. Zeng, Y. Xiong, F. Liu, J. Ye, and J. Tang, "Uncovering the spatiotemporal patterns of traffic congestion from large-scale trajectory data: A complex network approach," *Phys. A, Stat. Mech. Appl.*, vol. 604, Oct. 2022, Art. no. 127871.
- [40] Y. Xu, X. Cai, E. Wang, W. Liu, Y. Yang, and F. Yang, "Dynamic traffic correlations based spatio-temporal graph convolutional network for urban traffic prediction," *Inf. Sci.*, vol. 621, pp. 580–595, Apr. 2023.
- [41] C. F. Daganzo, *Fundamentals of Transportation and Traffic Operations*. Oxford, U.K.: Pergamon, 1997.
- [42] D. C. Montgomery, C. L. Jennings, and M. Kulahci, *Introduction to Time Series Analysis and Forecasting*. Hoboken, NJ, USA: Wiley, 2015.
- [43] E. Taskesen. (2020). *Distfit is a Python Library for Probability Density Fitting*. [Online]. Available: <https://erdogant.github.io/distfit>
- [44] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 1–37, 2011.
- [45] (2010). *December 2010 North American Blizzard*. Accessed: Oct. 6, 2023. [Online]. Available: https://en.wikipedia.org/wiki/December_2010_North_American_blizzard

- [46] X. Wu, M. Xu, J. Fang, and X. Wu, "A multi-attention tensor completion network for spatiotemporal traffic data imputation," *IEEE Internet Things J.*, vol. 9, no. 20, pp. 20203–20213, Oct. 2022.
- [47] M. Nouri, M. Reisi-Gahrooei, and M. Ilbeigi, "A method for granular traffic data imputation based on PARATUCK2," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2676, no. 10, pp. 220–230, Oct. 2022.
- [48] Y. Zhang, X. Wei, X. Zhang, Y. Hu, and B. Yin, "Self-attention graph convolution residual network for traffic data completion," *IEEE Trans. Big Data*, vol. 9, no. 2, pp. 528–541, Apr. 2023.



Mina Nouri received the B.S. and M.S. degrees in civil engineering from the University of Tehran, Tehran, Iran. She is currently pursuing the Ph.D. degree with the Department of Civil, Environmental and Ocean Engineering, Stevens Institute of Technology, Hoboken, NJ, USA. Her research interests include machine learning, big data analytics, and anomaly detection, with applications to real-world systems.



Elif Konyar received the B.S. and M.S. degrees in industrial engineering from Boğaziçi University, Istanbul, Turkey. She is currently pursuing the Ph.D. degree with the Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL, USA. Her research interests include big data analytics and tensor-based modeling to model and monitor complex systems with high-dimensional data. She is a member of the Institute for Operations Research and the Management Sciences (INFORMS).



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 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.