This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2017.2744419, IEEE Transactions on Visualization and Computer Graphics

Voila: Visual Anomaly Detection and Monitoring with Streaming Spatiotemporal Data

Nan Cao, Chaoguang Lin, Qiuhan Zhu, Yu-Ru Lin, Xian Teng, Xidao Wen

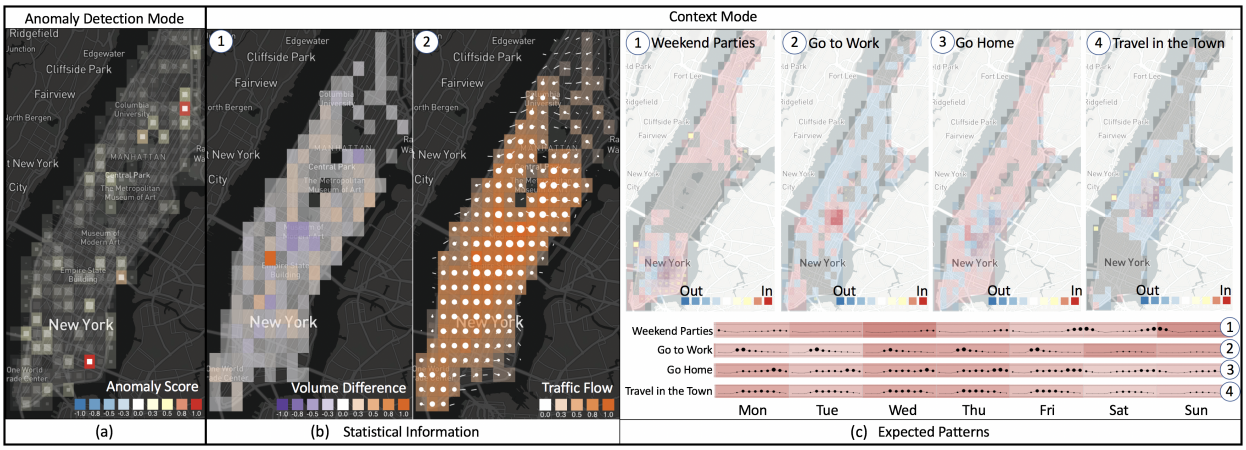

Fig. 1. Voila system employs a map visualization to provide an overview of the anomalous information in the form of a heatmap with visual cues to direct users' attention to the most "interesting" regions. Two different map modes are designed: the *anomaly detection mode* showing the regional anomaly scores and enabling an online anomaly inspection (a), and the *context mode* showing statistical context (b) and expected patterns (c). These modes and contexts can be switched by users.

Abstract— The increasing availability of spatiotemporal data continuously collected from various sources provides new opportunities for a timely understanding of the data in their spatial and temporal context. Finding abnormal patterns in such data poses significant challenges. Given that there is often no clear boundary between normal and abnormal patterns, existing solutions are limited in their capacity of identifying anomalies in large, dynamic and heterogeneous data, interpreting anomalies in their multifaceted, spatiotemporal context, and allowing users to provide feedback in the analysis loop. In this work, we introduce a unified visual interactive system and framework, Voila, for interactively detecting anomalies in spatiotemporal data collected from a streaming data source. The system is designed to meet two requirements in real-world applications, i.e., online monitoring and interactivity. We propose a novel tensor-based anomaly analysis algorithm with visualization and interaction design that dynamically produces contextualized, interpretable data summaries and allows for interactively ranking anomalous patterns based on user input. Using the "smart city" as an example scenario, we demonstrate the effectiveness of the proposed framework through quantitative evaluation and qualitative case studies.

Index Terms—Anomaly Detection, Visual Analysis

1 Introduction

The increasing availability of spatiotemporal data collected from various sources provides new opportunities for understanding the data in their spatial and temporal context. Finding abnormal patterns occurred at some locations and time is of particular interest in many applications such as aerology, public health surveillance, and urban computing. For example, urban scientists and analysts are interested in detecting sudden changes in traffic patterns in a city to help prevent future accidents [1, 5], likely through better traffic controls and route planning [6, 14, 20, 59, 60].

Traditionally, the problem of anomaly detection has been approached mainly through statistical and machine learning techniques [13, 21, 26]. However, the effectiveness of these techniques are often hinged by two major inherent challenges in anomaly detection problem: (1) there is often a lack of clear boundary between normal and abnormal cases, and (2) labeled data for training and verifying models are usually unavailable or difficult to collect. These challenges are further aggravated by the proliferation of big data where massive and continuous data flow in from various sources, likely in a streaming format, such as data constantly pulled from weather and traffic sensors, mobile devices and social media sites. On one hand, the various data inputs provide rich, spatiotemporal context information to inform the anomalous occurrences; on the other, the data exhibit very high veracity and volatility in so-called *normal* cases. The characteristics of big data stress the need for flexibly and adaptively identifying and interpreting normal and abnormal cases along with their rich context information. While visual analytics opens a new possibility to fulfill the need, existing solutions (e.g., [12, 33, 38, 49]) so far are not equipped to handle the large, complex and dynamic data environment.

Anomaly detection is one of the important information processing tasks where visual analytics can be advantageous. However, most existing solutions have overlooked two types of information that are *external* to the machine analytics and visual mapping processes: (a) the dynamic inputs that drive the changes in anomaly definition, and (b) the addi-

- *Nan Cao, Chaoguang Lin, Qiuhan Zhu are with Intelligent Big Data Visualization (iDV^x) Lab, Tongji University. E-mail: {nan.cao|chaoguang.lin|qiuhan.zhu}@tongji.edu.cn*
- *Yu-Ru Lin, Xian Teng, and Xidao Wen are with University of Pittsburgh. Yu-Ru Lin is the contact author. E-mail: {yurulin|xian.teng|xidao.wen}@pitt.edu*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2017.2744419, IEEE Transactions on Visualization and Computer Graphics

tional human knowledge that is either undefined or unavailable to the machine-centric approach, referred to as the "soft" knowledge [47]. We identify three technical challenges in anomaly detection with streaming, spatiotemporal data: (1) **Adaptivity** toward dynamic, rich context data: the spatiotemporal data are big, dynamic, heterogeneous, and multidimensional; capturing anomalous patterns in such data and at the same time adapting to the data changes and human knowledge accumulated in the system is beyond the capacity of the machine-centric, anomaly detection approach. (2) **Interpretability**: representing such data as well as anomalous patterns intuitively and comprehensively, along with their spatiotemporal context, is difficult with the off-the-shelf visualization solutions [12, 33, 38, 49] or with the modern spatiotemporal visualization platforms [2, 9, 14, 22, 55, 61]. (3) **Interactivity**: while there exist systems that tailor for monitoring spatiotemporal anomalies, the needs of supporting *online* anomaly investigation and incorporating human judgment to guide a system to produce better results has not been addressed.

In this paper, we introduce Voila (visual analysis of spatiotemporal data), a visual analytics system and framework for interactively detecting anomalies in spatiotemporal data collected from a streaming data source. Our work has the following key contributions:

- **System:** We formulate the system design requirements and propose an integrated, visual analytics system that simultaneously tackles adaptivity, interpretability, and interactivity challenges. The system comprises an online data processing pipeline that unceasingly connects streaming data input to the adaptive analysis, visualization, and interaction.
- **Algorithm:** We propose a novel tensor-based anomaly analysis algorithm that not only adapts to the dynamics in the input data but also produces descriptive patterns that can be visually presented along with their spatiotemporal context.
- **Visualization and Interaction:** We propose a set of novel visualization and interaction designs that support users' interpretability and interactivity – in particular, the information foraging and the sensemaking of normal and abnormal patterns. Moreover, we propose a unique interaction framework that enables the system to incorporate users' judgment with machine analytics to aid their information foraging based on a Bayesian approach.

2 Related Work

In this section, we review techniques that are most relevant to our work, including the anomaly detection algorithms, visual anomaly detection, and visualizations for the spatiotemporal data.

2.1 Anomaly Detection Algorithms

Anomaly detection, given its wide range of applications, has been extensively studied over the past decades [13, 21, 26]. Various techniques, including statistic-based methods [27, 44, 56], classification-based algorithms (either supervised [25, 37, 54] or semi-supervised [15, 35]), distance-based algorithms [7, 8, 17, 24, 40], and spectral-based algorithms [46, 48], have been proposed to tackle the problem in different situations. All the techniques discussed here are not exhaustive, rather representative of the different approaches (more comprehensive reviews can be found in [13, 26]). These techniques are useful in producing numeric results of anomalies, e.g., the outlierness scores, but are limited in offering *interpretation* of the anomalies – that is, what features and context exhibited in an abnormal case. Furthermore, most of these techniques lack the capacity to deal with the multi-way or multifaceted features, such as features over time *and* space.

The tensor-based methods have been recently proposed to deal with the multifaceted features [21]. These methods leverage tensor decomposition to produce compact feature descriptors along multiple facets, with the advantage to examine features associated with the original, rich context (such as time or space). Most of the tensor-based methods are supervised [4,43] or semi-supervised [32,39,50], which rely on training models based on labeled cases. Almost all these techniques assume relatively stable patterns in normal cases where the labeled data are easy to collect, and the decomposition can be done in a batch process. This assumption makes it difficult to adapt to the online or streaming situation where the normal patterns may change over time. Most importantly, the interpretability of the tensor decomposition results has not been examined. Recently, Fanaee-T and Gama [20] introduced an unsupervised approach to detect events from a traffic tensor; however, like other methods, the produced feature matrices and their interpretation in the spatiotemporal context have been overlooked.

There have been techniques specifically designed for detecting anomalies in spatiotemporal data for various purposes, e.g., monitoring the sensor networks [53], detecting anomalies in the spatial-temporal network data [41, 57], and finding the change of climate or environment [16, 18, 19]. These algorithms are designed specifically for their applications based on the domain-specific assumptions and knowledge. More general algorithms without domain-specific assumptions have also been recently introduced [36]. However, a crucial limitation of all these approaches is that there has not been a systematic and unified way that helps interpret the results derived from these techniques. In this work, we present a unified framework that allows for detecting and comprehending normal and abnormal cases situated in the streaming, spatiotemporal data scenarios.

2.2 Visual Anomaly Detection

As discussed earlier, the ability to interpret normal and abnormal cases are crucial because of the two major inherent challenges in anomaly detection problem: (1) the lack of clear boundary between normal and abnormal cases, and (2) the difficulty of obtaining labeled data for training and verifying models. As a result, human experts' domain knowledge and experience need to be involved in judging the cases. An intuitive, comprehensible visual representation of the data or analysis results is thus extremely useful for supporting interpretation and facilitating a better decision making. Novel visual design [10] and visual analysis systems for anomaly detection have been proposed recently, such as systems for detecting spreading of rumors [58] and anomalous user behaviors [11]. Mckenna et al. [38] studied a cybersecurity dashboard visualization that helps network analysts identify anomalous patterns. There have been visual analysis systems developed for detecting anomalies in spatiotemporal data [12, 33, 49], e.g., by employing conventional supervised algorithms [33]. All these systems, however, are restricted in the way human judgment can be used to *guide* the systems to produce comprehensible results more efficiently during the analysis process.

Different from all the aforementioned systems, we introduce a visual interactive framework with a novel tensor-based unsupervised algorithm. Our approach not only achieves an anomaly detection performance better than many existing algorithms but also produces comprehensible visual representations that allow human analysts to examine the cases within the rich spatiotemporal context. Moreover, we provide a unique, adaptive anomaly investigation mechanism, which incorporates human judgment to instantly guide the detection algorithm to produce a refined set of anomalies. The proposed framework allows for analyzing and monitoring spatiotemporal data in general big data scenarios – where data are gathered from multiple input streams and near-real-time analysis and visualization are desirable.

2.3 Visualizing the Spatiotemporal Data

Spatiotemporal visualization is an interdisciplinary topic that involves techniques from geographic information system (GIS), information visualization, and urban simulation and computing. Langran and Chrisman [31] introduced four primary ways to represent spatiotemporal data in GIS (space-time cubes, sequential snapshots, base-state with amendments, and space-time composites). Andrienko et al. [3] summarized characteristics of spatiotemporal data and further categorized the analysis tasks into the *elementary* tasks (e.g., given a moment, identifying spatial locations and objects) and the *general* tasks (e.g., comparing behaviors at the same or different time interval). These tasks have been used to guide the designs for many existing visual analysis systems, including the space-time cube based visualizations [30, 51, 52], visualizations for spatiotemporal analysis and modeling [2], and the systems designed for exploring and visualizing large trip data [22], traffic data [14], mobile phone data [9], and the telco-data [55, 61]. Some

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2017.2744419, IEEE Transactions on Visualization and Computer Graphics

of these systems are designed to reveal vehicle or human mobility patterns for urban planning purposes [60], including public transportation optimization [45] and optimal location selection [34].

Inspired by these early design guidelines and visualization designs, in this work, we propose a novel visual interactive framework, Voila, that is particularly suitable for tackling the analysis challenges in detecting and interpreting anomalies in big, spatiotemporal data. In our framework, the data are transformed into a sequence of tensor time series following both the space-time cube model and the sequential snapshot models. The proposed analysis algorithm then compares the current state of the data with the historical states (general tasks) to detect anomalies, and the visualization represents a snapshot of the data and enables a detailed exploration of anomalies given a particular time and user judgment (elementary task). We further consider the sense-making models for intelligence analysis [28, 42] to guide the design of our visual analytics components.

3 System Overview

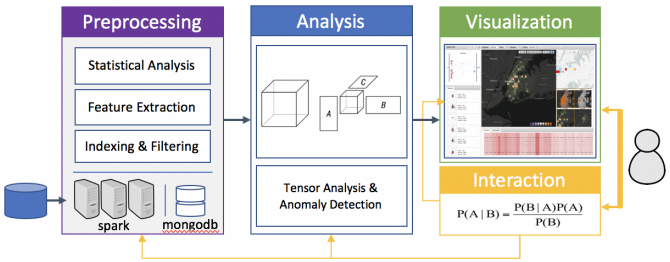

Fig. 2. The system architecture.

The proposed Voila system was designed to meet several real-world requirements for detecting anomalous patterns in the streaming spatiotemporal data. The design objective is to address the three technical challenges discussed earlier, whereas the concrete requirements are formulated through close collaboration with a domain expert whose expertise was in anomaly detection and spatiotemporal data analysis. Over the course of approximately nine months, regular meetings were held in which detailed system design requirements were discussed, and prototype systems were demonstrated to the expert for the purposes of gathering feedback. Three prototypes were built and improvements were made iteratively throughout the process. Below we describe the most critical design requirements (**R1**-**R3**) that were developed during these discussions, which motivate the design adopted in our work.

R1 **Adaptivity: Online monitoring and analysis.** The system should be efficient enough to perform a real-time or near real-time monitoring and analysis, given streaming data inputs, so that in the real-world applications the suspicious or abnormal cases detected from the data could be examined in a timely manner.

R2 **Interpretability: Multifaceted pattern discovery and anomaly filtering.** The system should create easy-to-understand designs to assist users in discovering abnormal patterns and help them understand "when and where, what might happen," with rich, spatiotemporal context. It should also direct users' attention to more significant anomaly instances.

R3 **Interactivity: Human in the analysis loop.** Users should be able to provide their judgment during the analysis and guide the system to produce the *refined* analysis results in real-time.

Based on these system requirements, we developed Voila, an integrated system solution for visual analytics addressing anomaly detection in spacial-temporal data. The system employs intuitive visualization designs as well as the anomaly detection techniques designed for multi-way structures to help identify suspicious data patterns in spatiotemporal data. Figure 2 illustrates the system architecture and the interactive analysis pipeline that it supports. The system transforms spatiotemporal data into tensor time series and derives and represents the abnormal patterns via four major modules: (1) the data preprocessing module, (2) the analysis module, (3) the visualization module, and (4) the interaction module. In particular, the data preprocessing module transforms the raw data into a series of multi-way tensors with each tensor representing multifaceted data in a given *time epoch* in which the anomalous patterns will be investigated. As illustrated in Fig. 3, new tensors capturing data in the later time epochs can be incrementally added into the series of tensors for analysis. This streaming pipeline facilitates the online monitoring and analysis (**R1**). In our implementation, the data preprocessing module runs in parallel on a spark cluster and the processed results are stored using the MongoDB[1] database for later retrieval of the results. The analysis module includes a novel, tensor-based anomaly detection algorithm that derives interpretable, anomalous patterns from streaming inputs (**R1,2**). The visualization module comprises a set of rich-context views that show the spatiotemporal patterns and reveal suspicious instances derived from the tensor analysis (**R2**). The interaction module provides an on-line, user feedback mechanism based on a Bayesian updating rule. With this module, users can guide the system to re-order the important anomalous patterns in real-time by incorporating users' judgment and observations without the re-computation of the tensor analysis, which makes the system salable and responsive to the user interactions (**R3**).

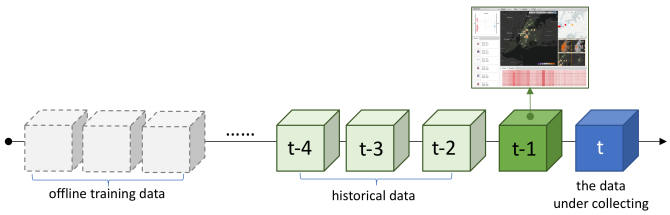

Fig. 3. The system pipeline for data processing and analysis.

As shown in Fig. 3, the system monitors the spatiotemporal data collected from a streaming data source in a near-real-time. We describe a system that runs in real-time if it is fast enough to calculate the data while collecting them, i.e., the computation delay can be ignored. We note that our system runs in near-real-time as the data analysis and the visualization of analysis results are lagging slightly behind the data collection process. In particular, within each time span in the time series, the system collects and transforms the data into a tensor, thus producing a tensor time series. When sufficient initial data are accumulated, the system continues analyzing and visualizing the data collected at past time epoch $(t-1)$, while keeping the new data arriving at the current time epoch (t). The historical data, whenever available, can be analyzed offline through an unsupervised learning procedure to enhance and calibrate the online analysis. The granularity of the time epoch can be chosen to be an hour, a day, a week, or a month, depending on the data input and the computational capacity of the data preprocessing and analysis modules. Details about the tensor-based data model and analysis methods will be discussed in the next section.

4 Anomalous Pattern Extraction

In this section, we introduce a novel tensor-based anomalous pattern discovery algorithm. Given a streaming input of spatiotemporal data, the objective of this algorithm is to identify any suspicious regions that potentially contain anomalies in an investigation space over time. After transforming the streaming spatiotemporal data into a time series of tensors, the algorithm produces a set of expected patterns based on the common distributions captured in the historical data, and identifies the regions in which the newly observed, empirical patterns are significantly deviated from what would be expected as anomalies. The analysis procedure comprises four key steps as shown in Fig. 4(1-4), which are detailed below.

4.1 Data Model and Transformation

The first step (Fig. 4(1)) is to transform the spatiotemporal data from a streaming data source (Fig. 4(a)) into a tensor time series as shown in Fig. 4(b).

A *tensor*, denoted as $\mathscr{X}$, is a mathematical representation of a multi-dimensional array, i.e., an extension of concepts such as scalars

[1] https://www.mongodb.com/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2017.2744419, IEEE Transactions on Visualization and Computer Graphics

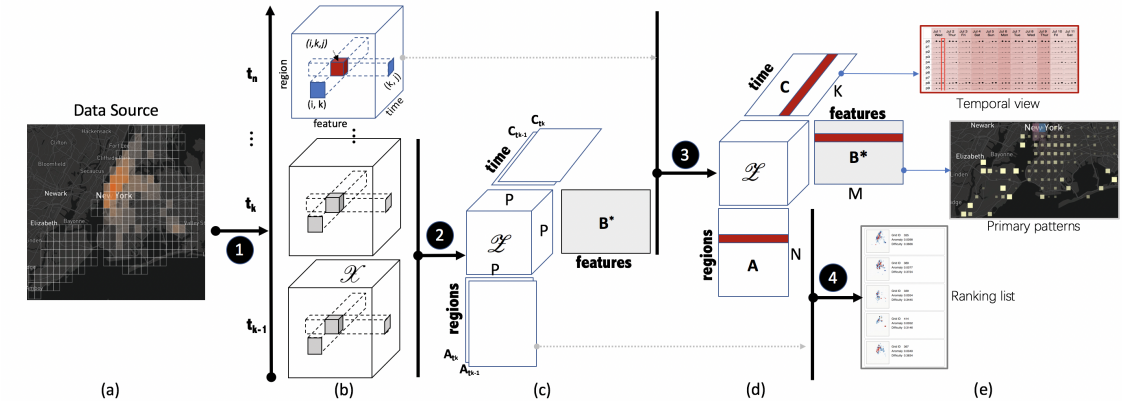

Fig. 4. The visual anomaly detection in the streaming spatiotemporal data consists of four major steps: (1) transforming the streaming spatiotemporal data into a tensor time series, (2) the expected pattern analysis based on historical data, (3) context analysis based on tensor decomposition, and (4) online regional anomaly detection in context.

(denoted as x), vectors (denoted as $\mathbf{x}$), and matrices (denoted as $\mathbf{X}$) to higher dimensions. A tensor is called n-way tensor if it has n-dimensions or *modes*. The dimensionality of each mode is determined by the number of its containing elements. For example, the three-way tensor $\mathscr{X} \in R_+^{I_1 \times I_2 \times I_3}$ has three modes with the dimensionality of I_1, I_2, and I_3, respectively. R_+ indicates that all the elements of $\mathscr{X}$ contain non-negative values, which commonly applies to situations when data represents numbers of observed instances.

The data that capture the spatiotemporal changes in various features can be modeled using the tensor representation – to reflect the "space-feature-time" associations. Given a set of K dimensional features, the data can be represented as a three-way tensor, denoted as $\mathscr{X} \in R_+^{N \times K \times M}$, where N indicates the number of regions in the space and M indicates the number of time steps in an epoch to be investigated. Here, the regions could be chosen as the administrative divisions in a city, or a set of grid cells that uniformly partition the geographical scope of a city. Each element of the tensor, $\mathscr{X}_{ikj}$, indicates the value of the k-th feature in the region i at time j.

Note that there are two notions of temporal resolution: the resolution of data collection, denoted by *time epoch*, and the resolution of analysis, denoted by *time step*. As shown in Fig. 3, the streaming spatiotemporal data are collected on the basis of time epoch. For data within a time epoch, we create a tensor in a finer granularity based on the resolution of time step.

We illustrate the tensor data representation using a traffic dataset as an example. The data can be collected and processed on a daily basis, results in a tensor $\mathscr{X}$ in each day capturing the changes of traffic features (Fig. 4(b)). In this scenario, the spatial granularity of $\mathscr{X}$, i.e., a region, could be a grid cell in a set of equal-size grid cells that uniformly partition the space (i.e., a city) where the data were collected (Fig. 4(a)). The temporal granularity of $\mathscr{X}$, i.e., a time step, could be an hour within the scope of a day (time epoch). We construct features that capture all traffic flows (number of vehicles) in the city within a day to be represented in a tensor. Specifically, for the i-th region at the j-th time step, we construct a $2N$-dimensional feature vector, denoted as $\mathscr{X}[i,:,j]$. The first N dimensions capture the numbers of outgoing traffic flow leaving from the i-th region to every other region, i.e., $\mathscr{X}[i,k,j]$ indicates the traffic flow from the i-th region to the k-th region at time j. The next N dimensions capture the incoming flow entering the i-th region from every other region, i.e., $\mathscr{X}[i,N+k,j]$ indicates the flow from the $(N+k)$-th region to the i-th region at time j. $\mathscr{X}[i,i,j]$ and $\mathscr{X}[i,N+i,j]$ are two special cases: both indicate the internal flows within the i-th grid (one of which is omitted in the analysis to avoid redundancy).

To better understand the representation, we now describe a real-world taxi trip dataset that will be used in our experiment.

Dataset. The New York City taxi-trip data[2] collected in 2014 and 2015. We extracted a subset of the taxi-trip data containing over 100 million taxi trips within the Manhattan area. We partitioned the area into grids of 311 equal-sized regions (each has an area of $2 \times 2\ km^2$). The data were first processed offline using parallel Spark clusters and were converted into a time series of tensors to simulate the streaming input, with time epoch set as a day, and the time step as two hours. We extracted 622 features for each region, which capture a region's in-coming and out-going taxi flows. Thus, the series contain over 720 tensors (spanning two years), each with the shape $311 \times 622 \times 12$. We employed this time series as the input of the Voila to simulate the real-time monitoring scenario.

[2] www.nyc.gov/html/tlc/html/about/tri_record_data.shtml

4.2 Extracting Expected Patterns

The second step (Fig. 4(2)) is to extract the expected patterns that commonly exist in the space based on the historical data given in the input tensor time series to facilitate capturing the anomalous patterns.

Given spatiotemporal data stored in a tensor, we derive the expected, latent patterns from the data through tensor decomposition. *Tensor decomposition* is an analysis that factorizes a tensor into a superdiagonal core tensor multiplied by a matrix along each mode. For example, given P, the number of desired latent patterns, the decomposition of a three-way tensor $\mathscr{X} \in R_+^{N \times K \times M}$ is to optimize the following objective function:

$$min \left\| \mathscr{X} - [\![\mathscr{Z}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \right\| \tag{1}$$

where $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are three factor matrices, and $\|\cdot\|$ denotes the Frobenius norm of a tensor (as well as a matrix or a vector). $[\![\cdot]\!]$ denotes a PARAFAC-like decomposition [23], where $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are respectively in the shapes of $N \times P$, $K \times P$, and $M \times P$, with latent patterns captured through a reduced dimensionality P, and $\mathscr{Z} \in R_+^{P \times P \times P}$ is the core tensor whose diagonal elements represent the relative strength of the corresponding patterns. In particular, each column in $\mathbf{A}$ and $\mathbf{C}$ represents the distribution of the latent patterns over different regions and time, respectively. Each row in $\mathbf{A}$ and $\mathbf{C}$ is a P-dimensional vector, denoted as $\mathbf{A}[i,:]$ or $\mathbf{C}[j,:]$, which respectively indicates the likelihoods of the patterns occurred in the region i at time j, referred to as the patterns' *occurrence-likelihoods*. $\mathbf{B}$ is the latent "feature-pattern" matrix which provides each latent pattern by a set of co-occurring features. The analysis is analogous to the topic modeling in which latent topics interpreted by a set of relevant keywords are extracted from a corpus. Given the three-way tensor in form of "region-feature-time", the regions at a particular time are analogous to the documents, and the features are analogous to the keywords in documents. Thus latent patterns derived from the tensor are analogous to the latent topics derived from the document corpus. Here, a pattern is captured by a distribution over a set of relevant features as a latent topic captured by a distribution over a set of relevant keywords. Moreover, our tensor-based analysis taking into account the richer context information in the data, such as the temporal dimension, in a unified manner.

We propose the following analysis model to detect latent patterns from a tensor time series:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2017.2744419, IEEE Transactions on Visualization and Computer Graphics

$$\mathbf{B}^* = arg\,min \sum_{t=1}^{n} (\| \mathscr{X}_t - [\![\mathscr{Z}; \mathbf{A}_t, \mathbf{B}, \mathbf{C}_t]\!] \| + \alpha \|\mathbf{A}_t - \mathbf{A}_{t-1}\| + \beta \|\mathbf{C}_t - \mathbf{C}_{t-1}\|) \quad \text{subject to: } \mathbf{B}^T\mathbf{B} = \mathbf{I}, \mathbf{B} \geq 0, \mathbf{A}_t \geq 0, \mathbf{C}_t \geq 0, t = 1, 2, \ldots, M \tag{2}$$

in which, we fix $\mathbf{B}$ during the analysis to capture the innate patterns that remain unchanged over time in the investigation space. The constraint $\mathbf{B}^T\mathbf{B} = \mathbf{I}$ ensures all the patterns captured by $\mathbf{B}$ (i.e., $\mathbf{B}$'s column vectors) are orthogonal to each other, thus making them unique and distinguishable. The factor matrices $\mathbf{A}_t$ and $\mathbf{C}_t$ are smoothed according to their preceding matrices (the second and third regularization terms) to reduce data noises. The degree of smoothness is controlled respectively by the parameter α and β. All the factor matrices are constrained to be non-negative, which facilitates the interpretability of occurrence-likelihood.

We solve the above optimization problem using block coordinate descent [29] in an offline procedure based on a tensor time series of the historical data. The resulting matrix $\mathbf{B}^*$ captures the expected patterns occurred in the given space (e.g., a city) during the given period of time, which is used as the basis for calculating the regional anomaly scores.

4.3 Extracting Dynamic Patterns

In the step of context analysis (Fig. 4(3)), the algorithm investigates the expected patterns' occurrence-likelihoods in space and time based on a testing tensor $\mathscr{X}_t$ in the series via the following tensor decomposition:

$$min \| \mathscr{X}_t - [\![\mathscr{Z}; \mathbf{A}_t, \mathbf{B}^*, \mathbf{C}_t]\!] \| \tag{3}$$

where the expected patterns, i.e., $\mathbf{B}^*$ are preserved. $\mathbf{A}_t$ and $\mathbf{C}_t$ are factor matrices that respectively represent how the patterns are distributed in regions and time in the scope given by $\mathscr{X}_t$. In particular, $\mathbf{A}_t$ indicates the patterns' occurrence-likelihoods in regions, which is latter used for calculating the regional anomaly score. $\mathbf{C}_t$ captures the time-varying context information as shown in Fig. 4(e).As the expected patterns, $\mathbf{B}^*$, are preserved, $\mathbf{A}_t$ and $\mathbf{C}_t$ may vary significantly from the normal case if the real patterns in the analysis scope, $\mathscr{X}_t$, are different from the expected ones. In particular, detecting how patterns are varying in regions based on $\mathbf{A}_t$ will in particular helpful in terms of locating an anomaly at a given time, which will be discussed next.

4.4 Extracting Anomalous Patterns

Based on the above analysis of expected and dynamics patterns, we derive measures to quantify the extent to which a region would have anomalous patterns. We calculate an *anomaly score* s_i for each region r_i at a given time t by examining the changes of the patterns' occurrence-likelihoods in r_i based on $\mathbf{A}_t[i,:]$ where each element $\mathbf{A}_t[i,k]$ indicates the likelihood of the k-th latent pattern occurring in r_i at time t. Then, by comparing $\mathbf{A}_t[i,:]$ with its historical values $\{\mathbf{A}_{t'}[i,:]\}_{t'<t}$, we can determine how much the pattern of r_i at time t deviates from what would be expected from the historical data. The deviation can be computed by leveraging anomaly detection methods such as local outlier factor (LOF) [8] and One-Class SVM [15]. In our experiment, we find LOF performs better as it is more robust against false positive cases.

5 Visualization and Interaction

In this section, we introduce the visualization and interaction components. We begin with the design consideration, followed by the technical details of each component.

5.1 Design Tasks

The design of the visual analytics modules was iteratively refined based on the discussions with our domain expert. In particular, we discussed and formulated the key challenges and system requirements for visually analyzing anomalies in the large streaming spatiotemporal data and further identified a set of key features to be supported and key information to be represented through the visualizations. We then reconsidered these requirements based on the sensemaking model for intelligence analysis [28, 42]. One of the most widely used models, proposed by Pirolli and Card [42], describes two iterative loops in the sensemaking process: the *information foraging* loop and the *sensemaking* loop. The former involves processes aimed at seeking information, searching and filtering it, and reading and extracting information, while the latter involves iterative development of a mental model (a conceptualization or schematic representation) that best fits the evidence. These processes are not necessarily sequential or discrete, but can be parallel [28]. Guided by these characteristics, we organize our design consideration into seven design tasks (**T1**-**T7**). The first three tasks aim to support information foraging loop – to augment users' information seeking through overview (**T1**), ranking (**T2**), and linking to the raw data (**T3**). The next three tasks aim to support sensemaking loop – to augment users' conceptualization of normal and abnormal cases through showing patterns in context (**T4**), comparing patterns (**T5**), and external memorization (**T6**). The final task (**T7**) is to incorporate users' additional judgment to enhance both information foraging and sensemaking loops. We summarize the design tasks as follows.

- **T1 Overview of the investigation scope.** Due to the complexity (big, dynamic, heterogeneous, and multidimensional) of the spatiotemporal data, it is critical to provide a visualization that clearly illustrates the spatial and temporal scopes in which the anomalies are analyzed to help narrow down the searching space.
- **T2 Dynamic visual ranking of suspicious regions.** Visualization should be designed to aid the searching and filtering of anomalous information through visually revealing the suspicious regions in some order such that users' information seeking effort can be directed to more suspicious regions.
- **T3 Efficient browsing of the raw data.** The raw information, for example, the taxi trips entering or leaving a region in the New York City, shows evidence of that may be used to confirm or disconfirm an anomaly case. Therefore, the visualization should enable an efficient mechanism to extract and explore the raw data.
- **T4 Interpreting anomaly in their spatiotemporal contexts.** A schematic representation of anomaly patterns helps users develop a mental model to organize various types of normal and abnormal cases, which improves the capacity of users to attend to more of the structure of organized evidence [42]. The interpretation of these patterns needs to be established through showing the patterns along with their spatial and temporal occurrences (context), including the relevant raw data and features, the statistics derived from the historical data, and the occurrence-likelihood of the expected patterns.
- **T5 Facilitating visual data comparisons and correlations.** The visualization and interaction should be designed to enable users to make comparisons over patterns or data instances, or provide visual cues for finding relevant patterns and instances, to help users support, disconfirm, or re-evaluate their findings of anomalies.
- **T6 External memorization of anomaly analysis.** Once a suspicious region is attended, or an anomaly is identified, the system should allow users to record their analysis statuses and results into snapshots to facilitate a later review or further examination.
- **T7 Updating analytics based on human judgment.** The system should provide an interaction mechanism that allows incorporating human belief or judgment that is missing in machine training [47], such that all the views can be dynamically updated based on the human judgment and consequently redirect users' attention to alternative patterns or instances.

5.2 User Interface

Guided by the above design tasks and the expert user feedback, we develop our visualization and interaction components. The UI of Voila system, as shown in Fig. 5, consists of key views corresponding to each of the tasks: (1) the *macro map view* shows the overview of the anomaly detection results and within the spatial context (**T1**); (2) the *micro map view* and (3) the *history view* respectively show the spatial and temporal context of a focal region and the associated raw information, as well as the temporal statistics from the historical data to help examine the anomaly cases (**T3**); (4) the *temporal pattern view* shows the temporal distributions of the expected patterns (derived from the factor matrix $\mathbf{C}$) over the current and preceding epochs (**T4,5**); (5) the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2017.2744419, IEEE Transactions on Visualization and Computer Graphics

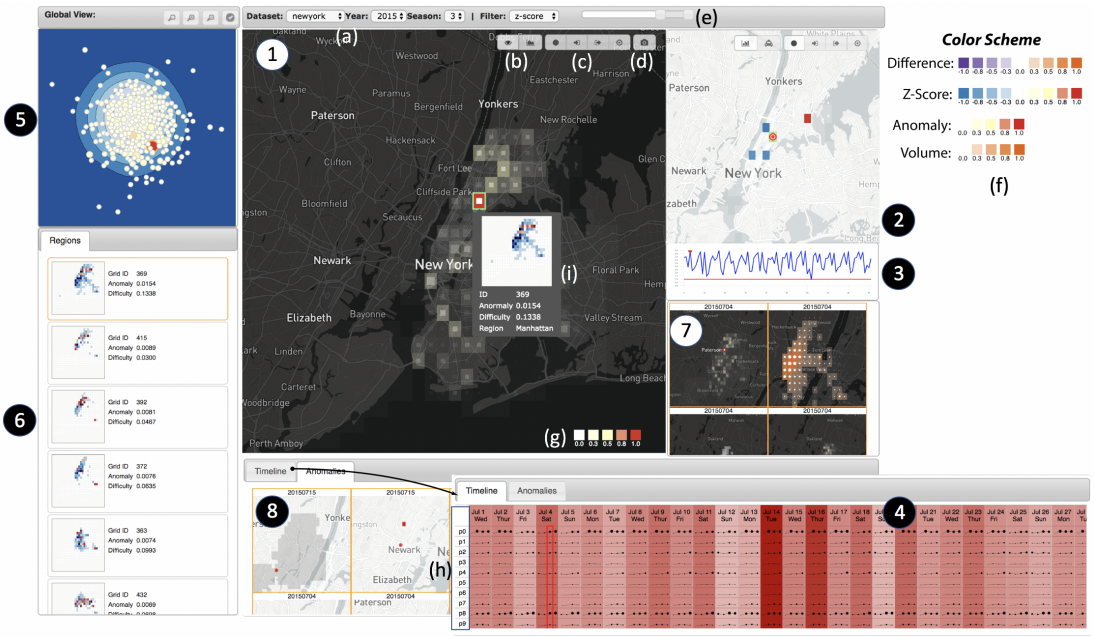

Fig. 5. The user interface of Voila system consists of eight major views: 1) macro and 2) micro map views; 3) history view; 4) temporal pattern view; 5) feature inspection view; 6) ranking list; 7) snapshot panel; and 8) anomaly panel.

snapshot panel and (6) the *anomaly panel* respectively allow users to record the suspicious regions and anomaly cases for later retrieval and further examination (**T6**); (7) the *ranking panel* orders the suspicious regions according to their anomaly scores updated based on the tensor analysis and user feedback (**T2**); (8) the *feature inspection view* allows for comparing how the regions at a given time are similar in terms of sharing similar features (**T5**). In all these views, regions are shown with colors encoding the anomaly-related information (detailed below) derived both from the tensor analysis and real-time user feedback (**T7**). Different color schemes are designed to illustrate different information such as the anomaly scores, the statistical information, the data difference, and the z-score values as shown in Fig. 5(f). To avoid ambiguity, an adaptive color legend is designed (Fig. 5(g)) which illustrates meaning and scale of colors shown in the corresponding views.

Usage scenario. To understand how Voila's different views work together, let us consider the following scenario. Take, for example, Mike, a security officer in the NYC Public Safety Department, whose duties are to ensure the safety and security in the city through routinely monitoring the city's traffic system, identifying anomalous traffic incidents, and being alert to potential hazard such as civil disorders. To increase his capacity, he uses the Voila system that takes streaming data from various sources (e.g. taxi trips, traffic sensors, etc.) as input. The system produces a near real-time overview of the traffic anomalies in the macro map view, showing a set of suspicious regions automatically. To see what happened in those regions, Mike first investigates the highly anomalous regions identified by the system; he picks such regions from the macro map view and ranking panel, checks the regions' raw trip information and statistics in the micro map view, and compares the information with historic data by navigating through the history view. When he confirms that certain regions are indeed anomalous, he clicks and marks those regions, and the system automatically captures those regions into the anomaly panel, updating all anomalous information based on Mike's input. Alternatively, if he finds suspicious regions that warrant further examination, he captures them into the snapshot panel. Later, he can retrieve these suspicious regions, compare them with other regions using feature inspection view, or explore different patterns typically seen in the city through temporal pattern view.

5.3 Macro Map View

The macro map view presents a geographical map overlaid with equal-size grids that uniformly partition the entire space to be investigated (e.g., a city). It provides an overview of the anomalous information in the form of a heatmap with visual cues to direct users attention to the most "interesting" regions. Two different modes are provided: (1) the *anomaly detection mode* (Fig. 1(a)) and (2) the *context mode* (Fig. 1(b,c)). Users can switch between the two modes based on the buttons as shown in Fig. 5(b,c).

5.3.1 Anomaly Detection Mode

The tensor-based anomaly detection algorithm identifies a set of suspicious regions with high anomaly scores (ref. Section 4). However, when a user (e.g., Mike) investigates these regions, s/he looks for evidence to support the anomaly identification. Therefore, users need to consider two types of complementary, contextual information when manually investigate certain regions: (1) how likely a region contains anomalies, and (2) how difficult it is to find out the anomaly evidence in the region. To assist users' inspection, the visual analysis system needs to help leverage the two types of information so users can decide where to pay their attention first, and how much efforts they should spend at different areas.

Context-guided inspection. To guide users' analysis with the aforementioned consideration, we associate each region i with two probabilistic quantities: p_i indicates the likelihood of the region i containing an anomaly, and q_i indicates the difficulty of finding an anomaly inside the region. Let s_i be the anomaly score of this region derived from the tensor analysis (Section 4.4). Then, p_i is given by normalizing the anomaly scores over the entire space:

$$p_i = \frac{s_i}{\sum_{j=0}^{N} s_j} \tag{4}$$

q_i should reflect the region's properties such as the volume of data and the number of road crossings that may result in the difficulties of identifying an anomaly. Take the taxi-trip data as an example, an anomalous event would be hard to detect if the region has many trips that enter from or head to diverse directions. Formally, let v_{ik} be the total number of trips from the region i to region k, and v_i is the total number of trips leaving (or entering) the region i. The *diversity* of the trip directions, d_i is estimated based on the information entropy:

$$d_i = -\sum_{k \neq i} \frac{v_{ik}}{v_i} log_2 \frac{v_{ik}}{v_i} \tag{5}$$

q_i is thus calculated by considering both the trip volume and the diversity of trip directions:

$$q_i = \sqrt{\frac{v_i}{\max_{0 \leq j \leq N} v_j} \cdot \frac{d_i}{\max_{0 \leq j \leq N} d_j}} \tag{6}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2017.2744419, IEEE Transactions on Visualization and Computer Graphics

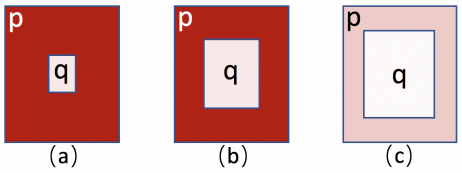

Fig. 6. The anomaly glyphs reveal both anomaly likelihood (saturation of the background color) and the difficulty of finding an anomaly (the size of the inner rectangle).

where q_i is normalized to $[0,1]$ as it represents probability.

Anomaly glyph. We design an *anomaly glyph* to show visual cues for the context-guided inspection. As shown in Fig. 6, an anomaly glyph for a region is represented as a rectangle corresponding to the boundary of the region, with the background color representing p, and the size of the inner rectangle representing $\sqrt{q}$, which visually enhances the value q that is usually too small to be visualized. For example, Fig. 6(a) illustrates the case in which the region is highly suspicious (with high p value shown in dark red in its background) but very easy to examine (with low q value shown with small inner rectangle). In comparison, Fig. 6(b) indicates the region is highly suspicious but difficult to examine and Fig. 6(c) indicates the region is not suspicious (light red in background), but if there is an anomaly, it is difficult to be found (large inner rectangle). Users usually only need to focus on those high suspicious ones with large p value.

Updating with user feedback. During the process of interactive anomaly inspection, users are guided to first inspect one of the most suspicious regions (one tha has the highest p value shown in an anomaly glyph with the darkest red color). After inspection, the user labels the region based on their judgments to indicate whether the region indeed contains an anomaly or not. To facilitate the succeeding analysis, once a region is labeled, the anomaly glyphs in the space should be always automatically updated to highlight the region to be investigated in the next. To achieve this goal, we propose a Bayesian approach to update regions' anomaly scores by incorporating human judgment in real time.

The Bayes' theorem calculates the probability of an event A given the prior knowledge of a certain condition B that is relevant to A, which is formally given as:

$$P(A|B) = P(B|A)P(A)/P(B) \tag{7}$$

In our problem, we estimate the probability $p'_j = P(A_j|B_k)$ of the event A_j ("the anomaly exists in the region j") occurred in condition of B_k (i.e., "the user fails to find the anomaly in the region k") at each inspection step. Specifically, when the user fails to find the anomaly in the region k, the probability of another region j containing an anomaly is calculated as:

$$p'_j = P(A_j|B_k) = \begin{cases} \dfrac{P(B_k|A_k)P(A_k)}{P(B_k)} = p_k \dfrac{q_k}{1-p_k(1-q_k)} & (j=k) \\ \dfrac{P(B_k|A_j)P(A_j)}{P(B_k)} = p_j \dfrac{1}{1-p_k(1-q_k)} & (j\neq k) \end{cases} \tag{8}$$

where $P(A_j) = p_j$ indicates the probability of an anomaly in region j. $P(B_k|A_j)$ is defined to be 1 for $k \neq j$ to reflect the user's belief that s/he cannot find an anomaly in the region k, regardless. $P(B_k|A_k) = q_k$ when $k = j$ as it reflects the probability of "failing to find an anomaly in the region i even there exist one due to certain difficulty." $P(B_k)$ indicates the probability that the user fails to find an anomaly in region k, which can be calculated in two conditions: $P(B_k|\overline{A}_k)P(\overline{A}_k)$ (the user fails to find an anomaly in the region k given no anomaly exist in it) and $P(B_k|A_k)P(A_k)$ (an anomaly exists in the region k, but the user fails to find it due to certain difficulty).

Formula 8 updates the anomaly scores of all the regions in the space when a region is inspected and labeled by the user. In particular, after the re-normalization such that $\sum_j p'_j = 1$, the probabilities of other regions $p'_j = p_k \cdot 1/(1-p_k(1-q_k)) > p_j$ increase (i.e., when $j \neq k$), whereas $p'_k < p_k$ decreases as the region k is just inspected and the user fails to find the anomaly in it, which is consistent with users' intuition.

5.3.2 Context mode

The context mode of the macro map shows the statistical information derived either from the raw data or the tensor analysis. These different statistics are shown with corresponding heatmaps as shown in Fig. 1(b). For example, when analyzing the NYC taxi-trip data, user can choose to show one of the primary statistic information associated with regions, including the number of incoming, out-going, and internal trips, as well as the differences between the number of incoming and out-going trips (Fig. 1(b-1)) per region. When showing the number of trips, each region is represented as a *flow glyph* with the direction given by the gradient of the statistics with respect to the region's neighborhood, which captures the potential traffic flow (Fig. 1(b-2)).

Another important context to be shown in the context mode is the expected patterns captured in the factor matrix $\mathbf{B}^*$ as described in Section 4.3. When analyzing the NYC taxi-trip data, $\mathbf{B}^* \in R_+^{2N\times P}$ captures P expected patterns via the in-coming and out-going trips through a concatenated $2N$-dimensional features, where N is the number of regions in the investigation space. In particular, the vectors $\mathbf{B}^*[0:N,j]$ and $\mathbf{B}^*[N+1:2N,j]$ respectively indicate the quantities of the in-coming and out-going trips in each region corresponding to the j-th pattern. This information can be visualized in a heatmap in the context view as shown in Fig.1(c) in which four different patterns are illustrated respectively. In particular, in this heatmap, a region's color is a blending between red and blue that respectively indicates the quantities of in-coming and out-going trips. The portions of blue (denoted as α) and red (denoted as β) during the blending is determined respectively by the portion of out-going and in-coming trips calculated as follows:

$$\alpha = \frac{\mathbf{B}^*[i,j]}{\mathbf{B}^*[i,j]+\mathbf{B}^*[N+i,j]}, \quad \beta = (1-\alpha) \tag{9}$$

5.4 Micro Map View and History View

When a user chooses to investigate a particular region, the region's detailed information, such as the raw data and relevant statistics, will be shown in the micro map view (Fig. 5(2)) and the history view (Fig. 5(3)). The micro map view depicts the relationship between the focal region and the remaining regions. A heatmap is generated on the map centered on the focal region, with colors reflecting the strength of the relationships. We show three types of relationships: (i) *raw* in-coming or out-going flows at a given time, directly extracted from the raw data, (ii) *expected* in-coming or out-going flows derived from the tensor analysis as described before (i.e., based on the $\mathbf{B}^*$ matrix), and (iii) *deviated* flow with respect to the focal region's historic data, which reveals how the flows at a given time deviate from their typical or normal values.

The history view shows the anomaly scores of the focal region over time as a time-series chart, from which a user can inspect how the focal region's abnormal behavior changes over time.

5.5 Temporal Pattern View

The temporal pattern view (Fig. 5(4)) visualizes the temporal distribution of the dynamic latent patterns derived from the tensor analysis, as described in Section 4.3. Specifically, the temporal factor matrix $\mathbf{C}_t$ of the current epoch t is visualized as a list of small multiple charts, where each row k in the temporal view represents the temporal distribution of the k-th latent pattern captured by the column vector $\mathbf{C}_t[:,k]$, with background color indicating the degree of anomaly for the corresponding time when compared with the history, which is determined by the largest regional anomaly score in the investigation space at the given epoch. While monitoring, the length of the temporal view increases when a new epoch is added into the system for investigation. The small multiple time series produced based on $\mathbf{C}_t$ of the current epoch are appended at the right end of the timeline.

5.6 Other Views

The system also provides other views for various purposes. In particular, **Feature Inspection View** (Fig. 5(5)) shows the similarity among regions in terms of sharing similar expected patterns. In this view, each region is represented as a circle with the color indicating its

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2017.2744419, IEEE Transactions on Visualization and Computer Graphics

anomaly score and size indicating the trip volume at the given time. The spatial distance between two regions on this view reflects their similarity which is calculated based on MDS projection. **Ranking Panel** (Fig. 5(6)) shows a list of highly suspicious regions ordered based on their anomaly scores. **Snapshot Panel** (Fig. 5(7)) shows the snapshots of the macro map view manually captured by the users during their inspection. **Anomaly Panel** (Fig. 5(8)) shows the snapshots of the micro map view automatically captured while the users click to inspect a suspicious region.

5.7 Interactions

The following interactions are designed to help with the data exploration and anomaly inspection. **Anomaly Inspection:** Users can left click on the macro-map view to select a region and show its context details in the micro map view and they can also single/double/ right click the region to label the cell as an abnormal/normal. The anomaly list will automatically captures the context of the anomaly shown in the micro map view into a picture. **Snapshot Capturing:** User can click the button shown in Fig. 5(d) to capture a snapshot of the current macro-map view into the snapshot view (Fig. 5(5)). **Context Switching:** Users can switch between different visualization modes to show different information via buttons shown in Fig. 5(b) (the anomaly detection mode), and Fig. 5(c) (the context mode). They can also to illustrate the patterns in the context mode of the macro map view by select a row in the temporal view via clicking the pattern labels(Fig. 5(h)). **Filtering:** In the system, a user can filter the regions shown in the macro and micro map views via a range slider (Fig. 5(e)) respectively based on their (1) anomaly scores, (2) the volumes of the containing data, and (3) z-scores. **Dynamic Zooming and Panning:** Both the macro-map view and the micro-map view support zooming and panning for exploring a large set of data items. Analyst can scroll the mouse to zoom and drag the mouse to pan. **Highlight, Link and Brush:** All the views are linked together. For example, when mouse hovers on a region in the map a tooltip showing the region's profile information and the z-map thumbnail centered at the region (Fig. 5(i)) will be displayed and all the corresponding regions in the list view, distribution view, and micro map view will be highlighted. Users can also brush to select a set of regions in the distribution view to highlight them in the macro mapview.

6 Evaluation

This section reports the evaluation of the proposed approach. We evaluated the algorithm effectiveness based on a quantitative evaluation, and studied the usability and usefulness of the system based on a case study with feedback from a domain expert. Both studies were conducted based on the New York City taxi-trip data described in Section 4.1.

6.1 Quantitative Evaluation

We evaluated the effectiveness of the proposed algorithm through a quantitative comparison with baseline methods based on human labeled ground truths.

Ground-truth labeling. We recruited 12 annotators to manually label the anomaly incidents within the Manhattan area over a six-month period (2014/01–2014/06). The annotators were asked to search online news reports exhaustively and any public information documented the event occurrence (e.g., the list of the NYC top events[3]). Over 300 potential anomaly events were first identified, and the annotators independently verified these events into a golden list of 96 anomaly incidents. Based on the list, the annotators manually label each region at each given time as positive if it corresponded to an anomaly occurrence, and negative otherwise.

Baseline methods and evaluation metrics. Two of the most widely adopted methods, *LOF* and *One-Class SVM* were chosen as the baseline. As described in Section 4.4, our algorithm augments both methods with tensor analysis. For brevity, we only report our best-performed method (tensor analysis with LOF). We compared our algorithm, denoted as "TA" (tensor-based anomaly detection) with the two baseline methods

[3]http://www.bizbash.com/new-yorks-top-100-events-2014/new-york/story/27977/

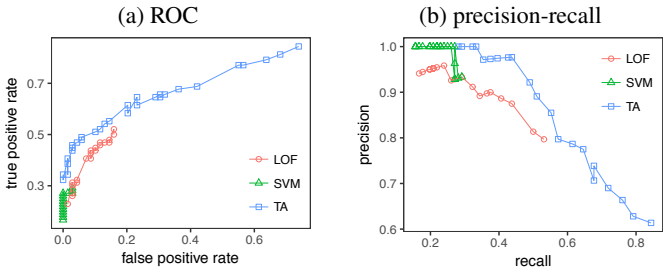

Fig. 7. Performance evaluation of anomaly detection. Results indicate our algorithm (TA) outperforms baseline methods (LOF and One-Class SVM).

in terms of standard information retrieval metrics: precision, recall, and ROC. As the ratio between positive and negative instances is extremely imbalanced, we use the precision-recall curves and ROC curves to discuss the results.

Evaluation results. As shown in Fig. 7, overall, our algorithm (TA) outperforms the baseline methods. The ROC plot (Fig. 7(a)) shows that, compared with LOF and One-Class SVM, TA had higher true positive rates when the false positive rates remain low (below 0.2). However, TA can achieve even higher positive rates while the baseline methods suffer from finding more true anomalies. (The One-Class SVM, for example, produced very low false positive rates but also could not reach high true positive rates.) The precision-recall plot (Fig. 7(b)) consistently reveals that TA can always achieve higher precision at the low recall conditions, but has the additional flexibility to achieve higher recall when slightly lower precision (e.g., 0.6–0.8) could be tolerated. This characteristic is crucial as it enables our system to retrieve a richer set of highly suspicious events that are otherwise difficult to find through the alternative algorithms. With the aid of system's visual analytics, users can then further investigate the set of suspicious events.

6.2 Case Study & Domain Expert Interview

We qualitatively studied the usability and usefulness of the Voila system through a semi-structured interview with a domain expert. The study serves two purposes: (a) to provide real examples to showcase the capability of the system, and (b) to provide user feedback on the system. Below, we describe the study set-up, the representative cases, and summarize the expert feedback on the system.

6.2.1 Study set-up and interview process

We invited an expert who has highly relevant expertise in public safety and traffic monitoring but has no prior knowledge about the incidences in our dataset. The expert is both an officer and a data scientist from the Institute of Public Safety in Shanghai (a city that has a scale comparable with the NYC). The interview lasted approximately 1.5 hours, which included two sessions: (1) introduction (30 minutes): an introduction to the dataset and the key features of the Voila system, followed by a tutorial and a user-practice sessions, and (2) discovery (60 minutes): a session where the expert was asked to use the system to identify anomalies, explore patterns, interpret and discuss her findings, and comment on the system capabilities. During the interview, a moderator was available to answer questions and record comments from the expert. The moderator also prompted questions for discussion, e.g., "How would you (or where did you) find that instance/pattern?" "How do you know this is an anomaly?" "What does it mean?" "How would this fit into your work?" "What worked well (or poorly) for you?" In the discovery session, the expert quickly identified several anomaly incidents and characteristic patterns. We briefly describe two representative cases in the following.

6.2.2 Representative cases

Exploring the expected patterns. The first feature that caught the expert's attention is the presentation of the city's expected patterns as well as the ability to browse these patterns. The expert explored the expected patterns by clicking each row of the temporal pattern view, and inspected the patterns on the macro map view using the context mode. The expert quickly found several interesting patterns (as

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2017.2744419, IEEE Transactions on Visualization and Computer Graphics

shown in Fig. 1(c)) where he can easily make sense. For example, a pattern (weekend parties; Fig. 1(c1)) captures the traffic flows within the downtown Manhattan (where the major arts/entertainment regions such as Broadway located) and between downtown and uptown (residential areas) that regularly occurred during the midnights in weekends. Other patterns capture the everyday routines in the city, such as the rush hour traffics in the mornings and evenings (Fig. 1(c2,c3), and the frequent travels in midtown in the afternoons (Fig. 1(c4)).

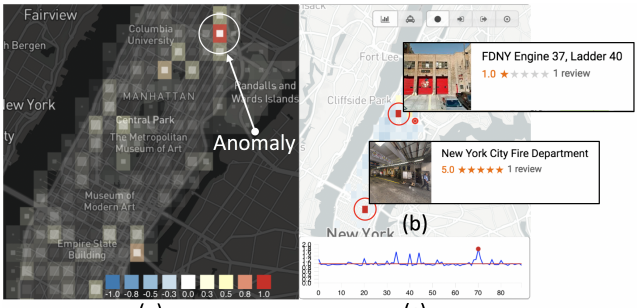

Fig. 8. The anonymous traffic flows during a gas explosion incident revealed in the system.

Detecting anomalies in Manhattan. After browsing through the expected patterns, the expert started engaging in finding and discussing the suspicious events. For example, when looking at the trips from 2014/03/12, a suspicious region in the uptown Manhattan, detected and highlighted by the system as shown in Fig. 8(a), caught the expert's attention. He clicked the region, trying to get the detailed spatial and temporal contexts from the the micro map view (Fig. 8(b)) and the history view (Fig. 8(c)). From the z-score based heatmap, he found anomalous taxi traffic flows between the focal region and two other regions – in particular, the z-scores are significantly higher compared with their historical values. He found one of them has the nearest fire station, and the New York City Fire Department is located in another region. Based on these, he conjectured a fire incident occurred in the focal region, which we verified as a gas explosion incident[4].

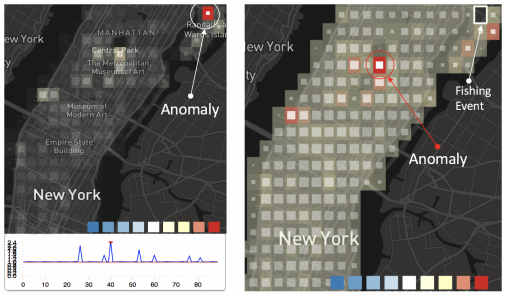

Fig. 9. Context-guided inspection reveals new anomalous region when receiving user feedback.

When investigating the potential anomalies, the expert also got very interested in Voila's context-guided inspection. For example, a suspicious region located in the Randalls and Wards Island was initially highlighted by the system (Fig. 9(a)). Based on the region's contextual information and the historical anomaly scores shown in the history view, the expert quickly found it was a park that frequently exhibited anomalous behaviors due to the frequent fluctuated traffic around the neighborhood, which we verified as a regular fishing contest[5]. He decided to mark the region as normal – based on this human judgment, the system instantly updated the anomaly scores of other regions. Then, another region was highlighted and immediately caught the expert's attention (Fig. 9(b)), which was verified as an irregular event (a public concert that attracted many people) happened on the same day.

[4]https://en.wikipedia.org/wiki/2014_East_Harlem_gas_explosion

[5]The fishing contests were frequently organized in that area, attracting a great number of people and resulting in anomalous traffic flows.

6.2.3 User feedback

During the discovery session, the expert provided a wealth of insightful feedback and comments, which we briefly summarize into four aspects. (1) **System:** The expert was impressed by the system design and the overall monitoring capability. "It's both cool and useful!" He commented, "especially for monitoring public safety," "we deal with the videos and images generated from the security cameras and video surveillance systems every day but have been struggled with finding important events from the enormous data; this system makes it possible to easily identify warning signs from big data." While he felt the information presented in the system is "very comprehensive," "it was a little overwhelming at first glance." He suggested that tacit tutorial or instructions could be added in the system to guide the first-time users. (2) **Visualization:** According to the expert's feedback, most of the visualization components successfully meet our design goals. For example, the expert felt the temporal pattern view and the heatmap visualization were fairly easy to comprehend and indeed helped discover many interesting patterns. "The anomaly glyph design is easy to understand; showing both anomaly scores and the difficulty of finding anomalies is very useful in real applications, although the updating rules seem to be complicated . . . " He commented, "the history view looks quite simple but very helpful in ruling out false positive cases." Compared with these views, "the MDS view seems less useful," and "the z-map requires a little more time to grasp the meaning." He particularly likes the visualization of expected patterns, "this is very interesting and informative . . . I cannot wait to see the patterns derived from Shanghai city!" (3) **Interaction:** He felt that showing the ranking of potential anomalies is "a new and smart idea." However, "when dealing with big data, it would be helpful if the system could tell users when to stop looking." Besides, the current system doesn't support "fact checking," and it took time for analysts to verify the detected anomalies, e.g., by searching events using a search engine; thus, he suggested, "it can be made even powerful to support the event search in the future." (4) **Usefulness and applicability:** The expert was confident that the system could be effectively used in smart city applications. He mentioned the stampede happened on New Year's Eve of 2015 in Shanghai, where at least 36 people died and 49 were seriously injured, "with this system, we could have seen the anomalous traffic flows a few hours before the event happened, and the accident might have been prevented." He further suggested that in the real city traffic monitoring scenarios, there is a need to use fine-grained temporal resolutions for data input (e.g., on a minute-by-minute basis), to incorporate as many data sources as possible (e.g., cellphone or other mobility data), and potentially to forecast or predict future incidents.

7 Conclusion and Future Work

In this paper, we introduce a visual analysis system, Voila, for interactively detecting anomalies in spatiotemporal data from a streaming source. The system is built based on several real-world requirements, such as online monitoring and interactivity. We propose a novel tensor-based anomaly detection algorithm with visualization and interaction design that dynamically produces contextualized, interpretable data summaries and allows for interactively ranking anomalous patterns based on user input. Using the NYC taxi-trip data, we evaluate the effectiveness and usefulness of Voila through a quantitative comparison and an expert interview. Our study results indicate the system's strengths and point out several new directions for future work, including providing tacit tutorials to guide the novice users, offering visual clues about low-precision instances in the anomaly ranking list, supporting fact search and checking, adaptively determining the temporal granularity, and developing new algorithms with forecasting and prediction capability.

Acknowledgments

We thank all the study participants and reviewers for their comments. This work is part of the research supported by NFSC Grants #61602306, the National Grants for the Thousand Young Talents in China, NSF Grants #1634944 and #1637067, and the CRDF & CIS at the University of Pittsburgh.

REFERENCES

[1] M. A. Abdel-Aty and A. E. Radwan. Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention*, 32(5):633–642, 2000.

[2] N. Andrienko and G. Andrienko. A visual analytics framework for spatio-temporal analysis and modelling. *Data Mining and Knowledge Discovery*, pp. 1–29, 2013.

[3] N. Andrienko, G. Andrienko, and P. Gatalsky. Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages & Computing*, 14(6):503–541, 2003.

[4] Y. Bai, J. Tezcan, Q. Cheng, and J. Cheng. A multiway model for predicting earthquake ground motion. In *International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, pp. 219–224. IEEE, 2013.

[5] T. C. Bailey and A. C. Gatrell. *Interactive spatial data analysis*, vol. 413. Longman Scientific & Technical Essex, 1995.

[6] M. Batty. The Size, Scale, and Shape of Cities. *Science*, 319(5864):769–771, Feb. 2008. doi: 10.1126/science.1151419

[7] S. D. Bay and M. Schwabacher. Mining distance-based outliers in near linear time with randomization and a simple pruning rule. In *Proceedings of SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 29–38. ACM, 2003.

[8] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander. Lof: identifying density-based local outliers. In *ACM SIGMOD Record*, vol. 29, pp. 93–104, 2000.

[9] F. Calabrese, L. Ferrari, and V. D. Blondel. Urban sensing using mobile phone network data: a survey of research. *Acm Computing Surveys*, 47(2):25, 2015.

[10] N. Cao, Y.-R. Lin, D. Gotz, and F. Du. Z-glyph: Visualizing outliers in multivariate data. *Information Visualization*, p. 1473871616686635, 2017.

[11] N. Cao, C. Shi, S. Lin, J. Lu, Y.-R. Lin, and C.-Y. Lin. Targetvue: Visual analysis of anomalous user behaviors in online communication systems. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):280–289, 2016.

[12] J. Chae, D. Thom, H. Bosch, Y. Jang, R. Maciejewski, D. S. Ebert, and T. Ertl. Spatiotemporal social media analytics for abnormal event detection and examination using seasonal-trend decomposition. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pp. 143–152, 2012.

[13] V. Chandola, A. Banerjee, and V. Kumar. Anomaly detection: A survey. *ACM Computing Surveys*, 41(3):15, 2009.

[14] W. Chen, F. Guo, and F.-Y. Wang. A survey of traffic data visualization. *IEEE Transactions on Intelligent Transportation Systems*, 16(6):2970–2984, 2015.

[15] Y. Chen, X. S. Zhou, and T. S. Huang. One-class svm for learning in image retrieval. In *International Conference on Image Processing*, vol. 1, pp. 34–37. IEEE, 2001.

[16] M. Das and S. Parthasarathy. Anomaly detection and spatio-temporal analysis of global climate system. In *Proceedings of International Workshop on Knowledge Discovery from Sensor Data*, pp. 142–150. ACM, 2009.

[17] T. De Vries, S. Chawla, and M. E. Houle. Finding local anomalies in very high dimensional space. In *International Conference on Data Mining*, pp. 128–137. IEEE, 2010.

[18] E. W. Dereszynski and T. G. Dietterich. Spatiotemporal models for data-anomaly detection in dynamic environmental monitoring campaigns. *ACM Transactions on Sensor Networks*, 8(1):3, 2011.

[19] J. H. Faghmous and V. Kumar. Spatio-temporal data mining for climate data: Advances, challenges, and opportunities. In *Data Mining and Knowledge Discovery for Big Data*, pp. 83–116. Springer, 2014.

[20] H. Fanaee-T and J. Gama. Event detection from traffic tensors: A hybrid model. *Neurocomputing*, 203:22–33, 2016.

[21] H. Fanaee-T and J. Gama. Tensor-based anomaly detection: An interdisciplinary survey. *Knowledge-Based Systems*, 98:130–147, 2016.

[22] N. Ferreira, J. Poco, H. T. Vo, J. Freire, and C. T. Silva. Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2149–2158, 2013.

[23] R. A. Harshman. Foundations of the parafac procedure: Models and conditions for an" explanatory" multi-modal factor analysis. 1970.

[24] V. Hautamaki, I. Karkkainen, and P. Franti. Outlier detection using k-nearest neighbour graph. In *Proceedings of International Conference on Pattern Recognition*, vol. 3, pp. 430–433. IEEE, 2004.

[25] S. Hawkins, H. He, G. Williams, and R. Baxter. Outlier detection using replicator neural networks. In *International Conference on Data Warehousing and Knowledge Discovery*, pp. 170–180. Springer, 2002.

[26] V. J. Hodge and J. Austin. A survey of outlier detection methodologies. *Artificial Intelligence Review*, 22(2):85–126, 2004.

[27] Z. Ju and H. Liu. Fuzzy gaussian mixture models. *Pattern Recognition*, 45(3):1146–1158, 2012.

[28] Y.-a. Kang and J. Stasko. Characterizing the intelligence analysis process: Informing visual analytics design through a longitudinal field study. In *IEEE Visual Analytics Science and Technology*, pp. 21–30, 2011.

[29] J. Kim, Y. He, and H. Park. Algorithms for nonnegative matrix and tensor factorizations: A unified view based on block coordinate descent framework. *Journal of Global Optimization*, 58(2):285–319, 2014.

[30] M.-J. Kraak. The space-time cube revisited from a geovisualization perspective. In *Proceeding of International Cartographic Conference*, pp. 1988–1996, 2003.

[31] G. Langran and N. R. Chrisman. A framework for temporal geographic information. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 25(3):1–14, 1988.

[32] J.-M. Lee, C. Yoo, and I.-B. Lee. On-line batch process monitoring using a consecutively updated multiway principal component analysis model. *Computers & Chemical Engineering*, 27(12):1903–1912, 2003.

[33] Z. Liao, Y. Yu, and B. Chen. Anomaly detection in gps data based on visual analytics. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pp. 51–58, 2010.

[34] D. Liu, D. Weng, Y. Li, J. Bao, Y. Zheng, H. Qu, and Y. Wu. Smartadp: Visual analytics of large-scale taxi trajectories for selecting billboard locations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):1–10, 2017.

[35] F. T. Liu, K. M. Ting, and Z.-H. Zhou. Isolation forest. In *International Conference on Data Mining*, pp. 413–422. IEEE, 2008.

[36] Y. Liu, B. Zhou, F. Chen, and D. W. Cheung. Graph topic scan statistic for spatial event detection. In *Proceedings of ACM International on Conference on Information and Knowledge Management*, pp. 489–498. ACM, 2016.

[37] M. V. Mahoney and P. K. Chan. Learning rules for anomaly detection of hostile network traffic. In *International Conference on Data Mining*, pp. 601–604. IEEE, 2003.

[38] S. McKenna, D. Staheli, C. Fulcher, and M. Meyer. Bubblenet: A cyber security dashboard for visualizing patterns. In *Computer Graphics Forum*, vol. 35, pp. 281–290. Wiley Online Library, 2016.

[39] P. Nomikos and J. F. MacGregor. Monitoring batch processes using multiway principal component analysis. *AIChE Journal*, 40(8):1361–1375, 1994.

[40] S. Papadimitriou, H. Kitagawa, P. B. Gibbons, and C. Faloutsos. Loci: Fast outlier detection using the local correlation integral. In *Proceedings of International Conference on Data Engineering*, pp. 315–326. IEEE, 2003.

[41] I. C. Paschalidis and G. Smaragdakis. Spatio-temporal network anomaly detection by assessing deviations of empirical measures. *IEEE/ACM Transactions on Networking (TON)*, 17(3):685–697, 2009.

[42] P. Pirolli and S. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of International Conference on Intelligence Analysis*, vol. 5, pp. 2–4, 2005.

[43] M. A. Prada, M. Dominguez, P. Barrientos, and S. Garcia. Dimensionality reduction for damage detection in engineering structures. *International Journal of Modern Physics B*, 26(25):1246004, 2012.

[44] P. J. Rousseeuw and A. M. Leroy. *Robust regression and outlier detection*, vol. 589. John wiley & sons, 2005.

[45] M. L. Sbodio, F. Calabrese, M. Berlingerio, R. Nair, F. Pinelli, et al. Allaboard: visual exploration of cellphone mobility data to optimise public transport. In *Proceedings of International Conference on Intelligent User Interfaces*, pp. 335–340. ACM, 2014.

[46] M.-L. Shyu, S.-C. Chen, K. Sarinnapakorn, and L. Chang. A novel anomaly detection scheme based on principal component classifier. Technical report, DTIC Document, 2003.

[47] G. K. L. Tam, V. Kothari, and M. Chen. An Analysis of Machine- and Human-Analytics in Classification. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):71–80, Jan. 2017. doi: 10.1109/TVCG.2016.2598829

[48] J. Terrell, K. Jeffay, F. D. Smith, L. Zhang, H. Shen, Z. Zhu, and A. Nobel. Multivariate svd analyses for network anomaly detection. In *Proceedings of ACM SIGCOMM Conference, Poster Session*, 2005.

[49] D. Thom, H. Bosch, S. Koch, M. Wörner, and T. Ertl. Spatiotemporal

anomaly detection through visual analysis of geolocated twitter messages. In *IEEE Symposium on Pacific Visualization*, pp. 41–48, 2012.

[50] X. Tian, X. Zhang, X. Deng, and S. Chen. Multiway kernel independent component analysis based on feature samples for batch process monitoring. *Neurocomputing*, 72(7):1584–1596, 2009.

[51] C. Tominski, P. Schulze-Wollgast, and H. Schumann. 3d information visualization for time dependent data on maps. In *Proceedings of International Conference on Information Visualization*, pp. 175–181. IEEE, 2005.

[52] U. D. Turdukulov, M.-J. Kraak, and C. A. Blok. Designing a visual environment for exploration of time series of remote sensing data: In search for convective clouds. *Computers & Graphics*, 31(3):370–379, 2007.

[53] X. R. Wang, J. T. Lizier, O. Obst, M. Prokopenko, and P. Wang. Spatiotemporal anomaly detection in gas monitoring sensor networks. In *Wireless Sensor Networks*, pp. 90–105. Springer, 2008.

[54] W.-K. Wong, A. Moore, G. Cooper, and M. Wagner. Bayesian network anomaly pattern detection for disease outbreaks. In *ICML*, pp. 808–815, 2003.

[55] W. Wu, J. Xu, H. Zeng, Y. Zheng, H. Qu, B. Ni, M. Yuan, and L. M. Ni. Telcovis: Visual exploration of co-occurrence in urban human mobility based on telco data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):935–944, 2016.

[56] K. Yamanishi, J.-I. Takeuchi, G. Williams, and P. Milne. On-line unsupervised outlier detection using finite mixtures with discounting learning algorithms. In *Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 320–324. ACM, 2000.

[57] W. C. Young, J. E. Blumenstock, E. B. Fox, and T. H. McCormick. Detecting and classifying anomalous behavior in spatiotemporal network data. In *Proceedings of KDD Workshop on Learning about Emergencies from Social Information (KDD-LESI 2014)*, pp. 29–33, 2014.

[58] J. Zhao, N. Cao, Z. Wen, Y. Song, Y.-R. Lin, and C. Collins. #fluxflow: Visual analysis of anomalous information spreading on social media. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1773–1782, 2014.

[59] Y. Zheng, L. Capra, O. Wolfson, and H. Yang. Urban Computing: Concepts, Methodologies, and Applications. *ACM Trans. Intell. Syst. Technol.*, 5(3):38:1–38:55, Sept. 2014. doi: 10.1145/2629592

[60] Y. Zheng, W. Wu, Y. Chen, H. Qu, and L. M. Ni. Visual analytics in urban computing: An overview. *IEEE Transactions on Big Data*, 2(3):276–296, 2016.

[61] Y. Zheng, W. Wu, H. Zeng, N. Cao, H. Qu, M. Yuan, J. Zeng, and L. M. Ni. Telcoflow: Visual exploration of collective behaviors based on telco data. In *IEEE International Conference on Big Data*, pp. 843–852. IEEE, 2016.