

Article

Spatiotemporal Data Mining: A Computational Perspective

Shashi Shekhar¹, Zhe Jiang^{1*}, Reem Y. Ali¹, Emre Eftelioglu¹, Xun Tang¹,
Venkata M. V. Gunturi² and Xun Zhou³

¹ Department of Computer Science and Engineering, University of Minnesota, Twin Cities. 4-192, Keller Hall, 200 Union St. SE, Minneapolis, MN 55455, USA;

E-Mails: shekhar@cs.umn.edu (S.S.); alireem@cs.umn.edu (R.Y.A.); emre@cs.umn.edu (E.E.); xuntang@cs.umn.edu (X.T.)

² Department of Computer Science and Engineering, Indraprastha Institute of Information Technology, Delhi. B-402 Academic Block, IIIT-Delhi Okhla, Phase III, New Delhi 110020, India;

E-Mail: gunturi@iiitd.ac.in (V.M.V.G.)

³ Department of Management Sciences, University of Iowa. S210 John Pappajohn Business Building, Iowa City, IA 52242, USA; E-Mail: xun-zhou@uiowa.edu (X.Z.)

* Author to whom correspondence should be addressed; E-Mail: zhe@cs.umn.edu;
Tel.: +1-612-626-7703.

Academic Editors: Emmanuel Stefanakis and Wolfgang Kainz

Received: 8 June 2015 / Accepted: 12 October 2015 / Published: 28 October 2015

Abstract: Explosive growth in geospatial and temporal data as well as the emergence of new technologies emphasize the need for automated discovery of spatiotemporal knowledge. Spatiotemporal data mining studies the process of discovering interesting and previously unknown, but potentially useful patterns from large spatiotemporal databases. It has broad application domains including ecology and environmental management, public safety, transportation, earth science, epidemiology, and climatology. The complexity of spatiotemporal data and intrinsic relationships limits the usefulness of conventional data science techniques for extracting spatiotemporal patterns. In this survey, we review recent computational techniques and tools in spatiotemporal data mining, focusing on several major pattern families: spatiotemporal outlier, spatiotemporal coupling and tele-coupling, spatiotemporal prediction, spatiotemporal partitioning and summarization, spatiotemporal hotspots, and change detection. Compared with other surveys in the literature, this paper emphasizes the statistical foundations of spatiotemporal data mining and provides comprehensive coverage of computational approaches for various pattern families.

We also list popular software tools for spatiotemporal data analysis. The survey concludes with a look at future research needs.

Keywords: spatiotemporal data mining; survey; review; spatiotemporal statistics; spatiotemporal patterns

1. Introduction

Explosive growth in geospatial and temporal data as well as the emergence of new technologies emphasize the need for automated discovery of spatiotemporal knowledge. Spatiotemporal data mining studies the process of discovering interesting and previously unknown, but potentially useful patterns from large spatial and spatiotemporal databases [1–5]. Figure 1 shows the process of spatiotemporal data mining. Given input spatiotemporal data, the first step is often preprocessing to correct noise, errors, and missing data and exploratory space-time analysis to understand the underlying spatiotemporal distributions. Then, an appropriate spatiotemporal data mining algorithm is selected to run on the preprocessed data, and produce output patterns. Common output pattern families include spatiotemporal outliers, associations and tele-couplings, predictive models, partitions and summarization, hotspots, as well as change patterns. Spatiotemporal data mining algorithms often have statistical foundations and integrate scalable computational techniques. Output patterns are post-processed and then interpreted by domain scientists to find novel insights and refine data mining algorithms when needed.

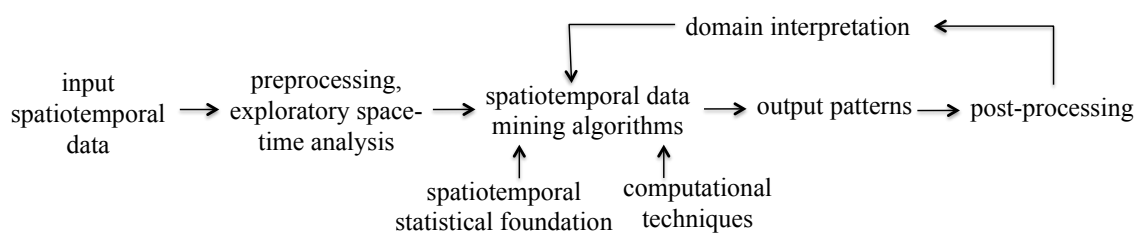


Figure 1. The process of spatiotemporal data mining.

Societal importance: Spatiotemporal data mining techniques are crucial to organizations which make decisions based on large spatial and spatiotemporal datasets, including NASA, the National Geospatial-Intelligence Agency [6], the National Cancer Institute [7], the US Department of Transportation [8], and the National Institute of Justice [9]. These organizations are spread across many application domains. In ecology and environmental management [10–13], researchers need tools to classify remote sensing images to map forest coverage. In public safety [14], crime analysts are interested in discovering hotspot patterns from crime event maps so as to effectively allocate police resources. In transportation [15], researchers analyze historical taxi GPS trajectories to recommend fast routes from places to places. Epidemiologists [16] use spatiotemporal data mining techniques to detect disease outbreak. There are also other application domains such as earth science [17], climatology [1,18], precision agriculture [19], and Internet of Things [20].

The interdisciplinary nature of spatiotemporal data mining means that its techniques must be developed with awareness of the underlying physics or theories in the application domains [21]. For example, climate science studies find that observable predictors for climate phenomena discovered by data science techniques can be misleading if they do not take into account climate models, locations, and seasons [22]. In this case, statistical significance testing is critically important in order to further validate or discard relationships mined from data.

Challenges: In addition to interdisciplinary challenges, spatiotemporal data mining also poses statistical and computational challenges. Extracting interesting and useful patterns from spatiotemporal datasets is more difficult than extracting corresponding patterns from traditional numeric and categorical data due to the complexity of spatiotemporal data types and relationships. According to Tobler's first law of geography, "Everything is related to everything else, but near things are more related than distant things." For example, people with similar characteristics, occupation and background tend to cluster together in the same neighborhoods. In spatial statistics such spatial dependence is called the spatial autocorrelation effect. Ignoring autocorrelation and assuming an identical and independent distribution (i.i.d.) when analyzing data with spatial and spatio-temporal characteristics may produce hypotheses or models that are inaccurate or inconsistent with the data set [23]. In addition to spatial dependence at nearby locations, phenomena of spatiotemporal tele-coupling also indicate long range spatial dependence such as El Niño and La Niña effects in the climate system. Another challenge comes from the fact that spatiotemporal datasets are embedded in continuous space and time, and thus many classical data mining techniques assuming discrete data (e.g., transactions in association rule mining) may not be effective. A third challenge is the spatial heterogeneity and temporal non-stationarity, *i.e.*, spatiotemporal temporal data samples do not follow an identical distribution across the entire space and over all time. Instead, different geographical regions and temporal period may have distinct distributions. Modifiable area unit problem (MAUP) or multi-scale effect is another challenge since results of spatial analysis depends on a choice of appropriate spatial and temporal scales. Finally, flow and movement and Lagrangian framework of reference in spatiotemporal networks pose challenges (e.g., directionality, anisotropy, *etc.*).

Previous surveys: As shown in Figure 2, surveys in spatial and spatiotemporal data mining can be categorized into two groups: ones without statistical foundations, and ones with a focus on statistical foundation. Among the surveys without focuses on statistical foundation, Koperski *et al.* [24] and Ester *et al.* [25] reviewed spatial data mining from a spatial database approach; Roddick *et al.* [12] provided a bibliography for spatial, temporal and spatiotemporal data mining; Miller *et al.* [26] cover a list of recent spatial and spatiotemporal data mining topics but without a systematic view of statistical foundation. Among surveys covering statistical foundations, Shekhar *et al.* 2003 [23] reviewed several spatial pattern families focusing on spatial data's unique characteristics; Kisilevich *et al.* [27] reviewed spatiotemporal clustering research; Aggarwal *et al.* [28] has a chapter summarizing spatial and spatiotemporal outlier detection techniques; Zhou *et al.* [29] reviewed spatial and spatiotemporal change detection research from an interdisciplinary view; Cheng *et al.* 2014 [30] reviews state of the art spatiotemporal data mining research including spatiotemporal autocorrelation, space-time forecasting and prediction, space-time clustering, as well as space-time visualization; Shekhar *et al.* 2011 [31] give the most recent review of spatial data mining research. However, its discussion of spatiotemporal patterns is limited (e.g., nothing on spatiotemporal change patterns or statistically significant hotspots).

In summary, there is no current survey in the literature that provides a systematic overview of spatiotemporal data mining that covers its statistical foundations as well as all major spatiotemporal pattern families.

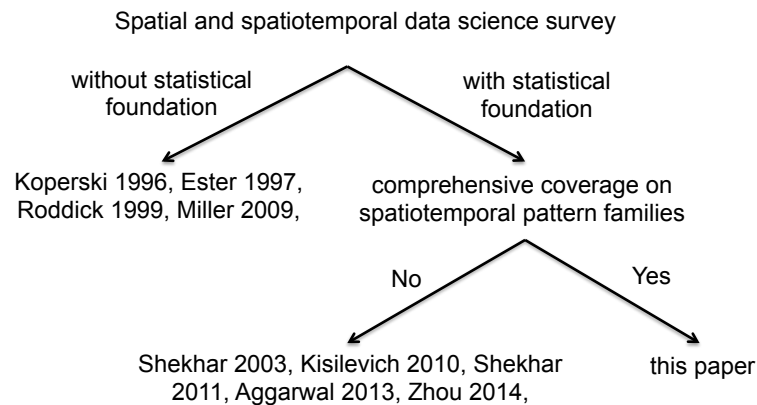


Figure 2. Categorization of spatial and spatiotemporal data mining surveys.

We hope this survey contributes to spatiotemporal data mining research in filling these two gaps. More specifically: (1) we provide a taxonomy of spatiotemporal data types; (2) we provide a taxonomy of spatial and spatiotemporal statistics organized by different data types; (3) we survey common computational techniques for all major spatiotemporal pattern families, including spatiotemporal outliers, spatiotemporal coupling and tele-coupling, spatiotemporal prediction, spatiotemporal partitioning and summarization, spatiotemporal hotspots and change patterns. Within each pattern family, techniques are categorized by the input spatiotemporal data types; (4) we analyze the research trends and future research needs.

Organization of the paper: This survey starts with the characteristics of the data inputs of spatiotemporal data mining (Section 2) and an overview of its statistical foundation (Section 3). It then describes in detail six main output patterns of spatiotemporal data mining related to outliers, association and tele-coupling, prediction, partitioning and summarization, hotspot, and change patterns (Section 4). Common software tools are discussed in Section 5. The paper concludes with an examination of research needs and future directions in Section 6.

2. Input: Spatial and Spatiotemporal Data

One important aspect of spatiotemporal data mining is its input data. This section provides a taxonomy of different spatial and spatiotemporal data types. The section also summarizes their unique data attributes and relationships. The goal is to provide a systematic overview of different techniques in spatiotemporal data mining tasks.

2.1. Types of Spatial and Spatiotemporal Data

The data inputs of spatiotemporal data mining tasks are more complex than the inputs of classical data science tasks because they include discrete representations of continuous space and time. Table 1 gives a taxonomy of different spatial and spatiotemporal data types (or models). Spatial data can

be categorized into three models, *i.e.*, the object model, the field model, and the spatial network model [3,32]. Spatiotemporal data, based on how temporal information is additionally modeled, can be categorized into three types, *i.e.*, temporal snapshot model, temporal change model, and event or process model [33–35]. In the temporal snapshot model, spatial layers of the same theme are time-stamped. For instance, if the spatial layers are points or multi-points, their temporal snapshots are trajectories of points or spatial time series (*i.e.*, variables observed at different times on fixed locations). Similarly, snapshots can represent trajectories of lines and polygons, raster time series, and spatiotemporal networks such as time expanded graphs (TEGs) and time aggregate graphs (TEGs) [36,37]. The temporal change model represents spatiotemporal data with a spatial layer at a given start time together with incremental changes occurring afterward. For instance, it can represent motion (e.g., Brownian motion, random walk [38]) as well as speed and acceleration on spatial points, as well as rotation and deformation on lines and polygons. Event and process models represent temporal information in terms of *events* or *processes*. One way to distinguish events from processes is that events are entities whose properties are possessed timelessly and therefore are not subject to change over time, whereas processes are entities that are subject to change over time (e.g., a process may be said to be accelerating or slowing down) [39].

Table 1. Taxonomy of Spatial and Spatiotemporal Data Models.

	Spatial Data	Temporal (Time Series)	Snapshots	Temporal Change (Delta/Derivative)	Events/Processes
object model	point(s)	1. point trajectories 2. spatial time series		displacement/motion (e.g., Brownian motion, random walk), speed/acceleration	spatial/spatiotemporal point process: Poisson, Cox, or Cluster process
	line(s)	line trajectories		motion/extension/rotation, deformation, split/merge	line process
	polygon(s)	polygon trajectories		motion/expansion/rotation/deformation, split/merge	flat process
field model	regular, irregular	raster time series		change across raster snapshots	cellular automation
spatial network model	graph	spatiotemporal network: 1. time expanded graph, time aggregated graph; 2. network flow		addition or removal of nodes and edges	1. random geometric graph; 2. spatiotemporal event or process on spatial network; 3. percolation theory

2.2. Data Attributes and Relationships

There are three distinct types of data attributes for spatiotemporal data: non-spatiotemporal attributes, spatial attributes, and temporal attributes. Non-spatiotemporal attributes are used to characterize non-contextual features of objects, such as name, population, and unemployment rate for a city. They are the same as the attributes used in the data inputs of classical data mining [40]. Spatial attributes are used to define the spatial location (e.g., longitude and latitude), spatial extent (e.g., area, perimeter) [41,42], shape, as well as elevation defined in a spatial reference frame. Temporal attributes include the timestamp

of a spatial object, a raster layer, or a spatial network snapshot, as well as the duration of a process. Relationships on these data attributes are summarized in Tables 2 and 3.

Table 2. Common Relationships among Non-spatial and Spatial Attributes.

Attributes	Categories	Relationships
non-spatial	<ul style="list-style-type: none"> • nominal • ordinal • interval • ratio 	Explicit <ul style="list-style-type: none"> • arithmetic • ordering • instance of • subclass of
spatial	<ul style="list-style-type: none"> • location • area • perimeter • shape 	Often implicit <ul style="list-style-type: none"> • set space: union, intersection, membership, <i>etc.</i> • topological space: meet, within, overlap, <i>etc.</i> • metric space: metric: distance, area, perimeter. directional: above, below, northeastern • others: shape based and visibility

Table 3. Relationships on Spatiotemporal Data.

Spatial Data		Temporal Snapshots (Time Series)	Change (Delta/Derivative)	Event/Process
object model	point(s),	1. spatiotemporal predicates [43]	1. displacement/motion speed/acceleration	1. spatiotemporal co-variance [38]
	line(s),	2. trajectory distance [44,45]	attraction/repulsion	2. spatiotemporal coupling for point events
	polygon(s)	3. spatial time series correlation [46], tele-connection [47]	2. (for line and polygon) extension/expansion, rotation, deformation, split/merge	or extended spatial objects [48–53]
field model	regular	1. cubic map algebra [54]	local, focal, zonal change	cellular automation [55]
	irregular	2. temporal correlation, tele-connection	across snapshots [29]	
spatial network	graph	1. predecessor/successor on a Lagrangian path 2. temporal centrality [56] 3. network flow [57–59]	change in centrality, connectivity	spatiotemporal coupling of network events

One way to deal with implicit spatiotemporal relationships is to materialize the relationships into traditional data input columns and then apply classical data mining techniques [60–64]. However, the materialization can result in loss of information [23]. The spatial and temporal vagueness which naturally exists in data and relationships usually creates further modeling and processing difficulty in spatial and spatiotemporal data mining. A more preferable way to capture implicit spatial and spatiotemporal relationships is to develop statistics and techniques to incorporate spatial and temporal information into the data mining process. These statistics and techniques are the main focus the survey.

3. Statistical Foundations

This section provides a taxonomy of common statistical concepts for different spatial and spatiotemporal data types. Spatial and spatiotemporal statistics are distinct from classical statistics due to the unique characteristics of space and time. One important property of spatial data is spatial dependency, a property so fundamental that geographers have elevated it to the status of the first law of geography: “Everything is related to everything else, but nearby things are more related than distant things” [65]. Spatial dependency is also measured using spatial autocorrelation. Other important properties include spatial heterogeneity, temporal autocorrelation and non-stationarity, as well as the multiple scale effect.

3.1. Spatial Statistics for Different Types of Spatial Data

Spatial statistics [38,66–68] is a branch of statistics concerned with the analysis and modeling of spatial data. The main difference between spatial statistics and classical statistics is that spatial data often fails to meet the assumption of an identical and independent distribution (i.i.d.). As summarized in Table 4, spatial statistics can be categorized according to their underlying spatial data type: Geostatistics for point referenced data, lattice statistics for areal data, and spatial point process for spatial point patterns.

Geostatistics: Geostatistics [69] deals with the analysis of spatial continuity (*i.e.*, dependence across locations), weak stationarity (*i.e.*, some statistical properties not changing with locations) and isotropy (*i.e.*, uniformity in all directions), which are inherent characteristics of spatial point reference data. It is mostly concerned with developing models of spatial dependence for prediction. Under the assumption of weak stationarity or intrinsic stationarity, spatial dependence at various distances can be captured by a covariance function or a semivariogram [66]. Geostatistics provides a set of statistical tools, such as Kriging [66] which can be used to interpolate attributes at unsampled locations. However, real world spatial data often shows inherent variation in measurements of a relationship over space, due to influences of spatial context on the nature of spatial relationships. For example, human behavior can vary intrinsically over space (e.g., differing cultures). Different jurisdictions tend to produce different laws (e.g., speed limit differences between Minnesota and Wisconsin). This effect is called spatial heterogeneity or non-stationarity. Special models (e.g., local space-time Kriging [70]) can be further used to reflect the varying functions at different locations.

Lattice statistics: Lattice statistics studies statistics for spatial data in the field (or areal) model. Here a lattice refers to a countable collection of regular or irregular cells in a spatial framework. The range of spatial dependency among cells is reflected by a neighborhood relationship, which can be represented by a contiguity matrix called a W-matrix. A spatial neighborhood relationship can be defined based on spatial adjacency (e.g., rook or queen neighborhoods) or Euclidean distance, or in more general models, cliques and hypergraphs [71]. Based on a W-matrix, spatial autocorrelation statistics can be defined to measure the correlation of a non-spatial attribute across neighboring locations. Common spatial autocorrelation statistics include Moran’s I , Getis-Ord G_i^* , Geary’s C , Gamma index Γ [69], *etc.*, as well as their local versions called local indicators of spatial association (LISA) [72]. Several spatial statistical models, including the spatial autoregressive model (SAR), conditional autoregressive model (CAR), Markov random fields (MRF), as well as other Bayesian hierarchical models [66], can be

used to model lattice data. Another important issue is the modifiable areal unit problem (MAUP) (also called the multi-scale effect) [73], an effect in spatial analysis that results for the same analysis method will change on different aggregation scales. For example, analysis using data aggregated by states will differ from analysis using data at individual family level.

Table 4. Taxonomy of Spatial and Spatiotemporal Statistics.

Spatial Model		Spatial Statistics	Spatiotemporal Statistics
object model	point(s)	Geostatistics (point reference data) 1. stationarity, isotropy, variograms, Kriging 2. local variogram or local correlation 3. the Modifiable Areal Unit Problem (MAUP)	Statistics for spatial time series 1. spatiotemporal stationarity, variograms, covariance, Kriging 2. temporal autocorrelation, tele-coupling 3. physics-inspired models 4. Dynamic Spatiotemporal model (Kalman filter)
	line(s)	line process	Spatiotemporal Point Processes 1. spatiotemporal (homogeneous) Poission process, Cox process, cluster process 2. spatiotemporal scan statistics 3. spatiotemporal K-function 4. Planetary/Brownian motion, random walk.
field model	regular, irregular	Lattice Statistics (areal data model) 1. W-matrix, spatial autocorrelation, local indicators of spatial association (LISA) 2. MRF, SAR, CAR, Bayesian hierarchical model 3. The Modifiable Areal Unit Problem (MAUP)	Statistics for raster time series 1. EOF analysis, CCA analysis 2. Spatiotemporal autoregressive model (STAR), Bayesian hierarchical model 3. Dynamic Spatiotemporal model (Kalman filter), data assimilation
	graph	Spatial Network Statistics 1. spatial network autocorrelation 2. network K function 3. network Kriging	

Spatial point processes: A spatial point process is a model for the spatial distribution of the points in a point pattern. It differs from point reference data in that the random variables are locations. Examples include positions of trees in a forest and locations of bird habitats in a wetland. One basic type of point process is a homogeneous spatial Poisson point process (also called complete spatial randomness, or CSR) [38], where point locations are mutually independent with the same intensity over space. However, real world spatial point processes often show either spatial aggregation (clustering) or spatial inhibition instead of complete spatial independence as in CSR. Spatial statistics such as Ripley’s K function [74,75], *i.e.*, the average number of points within a certain distance of a given point over the total average intensity, can be used to test a point pattern against CSR. Moreover, real world spatial point processes such as crime

events often contain hotspot areas instead of following homogeneous intensity across space. A spatial scan statistic [76] can be used to detect these hotspot patterns. It tests if the intensity of points inside a scanning window is significantly higher (or lower) than outside. Though both the K-function and spatial scan statistics have the same null hypothesis of CSR, their alternative hypotheses are quite different: the K-function tests if points exhibit spatial aggregation or inhibition instead of independence, while spatial scan statistics assume that points are independent and test if a hotspot with much higher intensity exists. Finally, there are other spatial point processes such as the Cox process, in which the intensity function itself is a random function over space, as well as a cluster process, which extends a basic point process with a small cluster centered on each original point [38]. For extended spatial objects such as lines and polygons, spatial point processes can be generalized to line processes and flat processes in stochastic geometry [77].

Spatial network statistics: Most spatial statistics research focuses on the Euclidean space. Spatial statistics on the network space is much less studied. Spatial network space, e.g., river networks and street networks, is important in applications of environmental science and public safety analysis. However, it poses unique challenges including directionality and anisotropy of spatial dependency, connectivity, as well as high computational cost. Statistical properties of random fields on a network are summarized in [78]. Recently, several spatial statistics, such as spatial autocorrelation, K-function, and Kriging, have been generalized to spatial networks [79–81]. Little research has been done on spatiotemporal statistics on the network space.

3.2. Spatiotemporal Statistics

Spatiotemporal statistics [38,82] combine spatial statistics with temporal statistics (time series analysis [83], dynamic models [82]). Table 4 summarizes common statistics for different spatiotemporal data types, including spatial time series, spatiotemporal point process, and time series of lattice (areal) data.

Spatial time series: Spatial statistics for point reference data have been generalized for spatiotemporal data [84]. Examples include spatiotemporal stationarity, spatiotemporal covariance, spatiotemporal variograms, and spatiotemporal Kriging [38,82]. There is also temporal autocorrelation and tele-coupling (high correlation across spatial time series at a long distance). Methods to model spatiotemporal process include physics inspired models (e.g., stochastically differential equations) [38] and hierarchical dynamic spatiotemporal models (e.g., Kalman filtering) for data assimilation [38].

Spatiotemporal point process: A spatiotemporal point process generalizes the spatial point process by incorporating the factor of time. As with spatial point processes, there is a spatiotemporal Poisson process, Cox process, and cluster process. There are also corresponding statistical tests including a spatiotemporal K function and spatiotemporal scan statistics [38].

Time series of lattice (areal) data: Similar to lattice statistics, there is spatial and temporal autocorrelation, a SpatioTemporal Autoregressive Regression (STAR) model [85], and Bayesian hierarchical models [66]. Other spatiotemporal statistics include empirical orthogonal functions (EOF) analysis (principle component analysis in geophysics), canonical-correlation analysis (CCA), and dynamic spatiotemporal models (Kalman filter) for data assimilation [82].

4. Output Pattern Families

4.1. Spatiotemporal Outlier

4.1.1. What are Spatiotemporal Outliers?

To understand the meaning of spatiotemporal outliers, it is useful first to consider global outliers. Global outliers [86–88] have been informally defined as observations in a data set which appear to be inconsistent with the remainder of that set of data, or which deviate so much from other observations as to arouse suspicions that they were generated by a different mechanism. In contrast, a spatiotemporal outlier [89–92] is a spatially and temporally referenced object whose non-spatiotemporal attribute values differ significantly from those of other objects in its spatiotemporal neighborhood. Informally, a spatiotemporal outlier is a local *instability* or *discontinuity*.

4.1.2. Application Domains

Detecting spatiotemporal outliers is useful in many applications including transportation, ecology, homeland security, public health, climatology, and location-based services [93,94]. For example, spatiotemporal outlier detection can be used to detect anomalous traffic patterns from sensor observations on a highway road network.

4.1.3. Statistical Foundation

The spatial statistics for spatial outlier detection are also applicable to spatiotemporal outliers as long as spatiotemporal neighborhoods are well-defined. The literature provides two kinds of bi-partite multidimensional tests: graphical tests, including variogram clouds [95] and Moran scatterplots [68,96], and quantitative tests, including scatterplot [97] and neighborhood spatial statistics [93,98].

4.1.4. Common Approaches

The intuition behind spatiotemporal outlier detection is that they reflect “discontinuity” on non-spatiotemporal attributes within a spatiotemporal neighborhood. Approaches can be summarized according to the input data types.

Outliers in spatial time series: For spatial time series (on point reference data, raster data, as well as graph data), basic spatial outlier detection methods, such as visualization based approaches and neighborhood based approaches, can be generalized with a definition of spatiotemporal neighborhoods. Thus, for simplicity, we only discuss these basic spatial outlier detection approaches here. The *visualization approach* plots spatial locations on a graph to identify spatial outliers. The common methods are variogram clouds and Moran scatterplot as introduced earlier. The *neighborhood approach* defines a spatial or spatiotemporal neighborhood, and a spatial statistic is computed as the difference between the non-spatial attribute of the current location and that of the neighborhood aggregate [93,99–101]. Spatial neighborhoods can be identified by distances on spatial attributes (e.g.,

K nearest neighbors), or by graph connectivity (e.g., locations on road networks) [89]. This research has been extended in a number of ways to allow for multiple non-spatial attributes [100], average and median attribute value [99,102], weighted spatial outliers [103], categorical spatial outlier [104], local spatial outliers [105], and fast detection [106,107].

Flow Anomalies: Given a set of observations across multiple spatial locations on a spatial network flow, flow anomaly discovery aims to identify dominant time intervals where the fraction of time instants of significantly mis-matched sensor readings exceeds the given percentage-threshold. Figure 3a is a simple example of problem input, which consists of two neighboring locations (*i.e.*, an upstream (up) and downstream (down) sensor), 10 time instants, and the notion of travel time (TT) or flow between the locations. The output contains two flow anomalies; using the time instants at the upstream sensor, periods 1–3 and 6–9, where the majority of time-points show significant differences in-between (Figure 3b). Flow anomaly discovery can be considered as detecting *discontinuities* or *inconsistencies* of a non-spatiotemporal attribute within a neighborhood defined by the flow between nodes, and such discontinuities are persistent over a period of time. A time-scalable technique called SWEET (Smart Window Enumeration and Evaluation of persistent-Thresholds) was proposed [57,108,109] that utilizes several algebraic properties in the flow anomaly problem to discover these patterns efficiently. To account for flow anomalies across multiple locations, recent work [58] defines a *teleconnected flow anomaly* pattern and proposes a RAD (Relationship Analysis of Dynamic-neighborhoods) technique to efficiently identify this pattern.

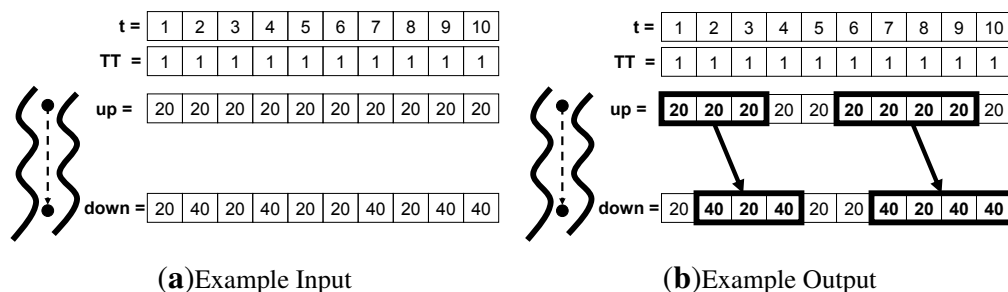


Figure 3. Flow Anomaly Example. (a) Example input. (b) Example output.

Anomalous moving object trajectories: Detecting spatiotemporal outliers from moving object trajectories is challenging due to the high dimensionality of trajectories and the dynamic nature. A context-aware stochastic model has been proposed to detect anomalous moving pattern in indoor device trajectories [110]. Another spatial deviations (distance) based method has been proposed for anomaly monitoring over moving object trajectory stream [111]. In this case, anomalies are defined as rare patterns with big spatial deviations from normal trajectories in a certain temporal interval. A supervised approach called Motion-Alert has also been proposed to detect anomaly in massive moving objects [112]. This approach first extracts motif features from moving object trajectories, then clusters the features, and learns a supervised model to classify whether a trajectory is an anomaly. Other techniques have been proposed to detect anomalous driving patterns from taxi GPS trajectories [113–115].

4.2. Spatiotemporal Couplings and Tele-Couplings

4.2.1. What are Spatiotemporal Couplings and Tele-Couplings?

Spatiotemporal coupling patterns represent spatiotemporal object types whose instances often occur in close geographic and temporal proximity. These patterns can be categorized according to whether there exists temporal ordering of object types: spatiotemporal (mixed drove) co-occurrences [48] are used for unordered patterns, spatiotemporal cascades [51] for partially ordered patterns, and spatiotemporal sequential patterns [53] for totally ordered patterns. Spatiotemporal tele-coupling [46] is the pattern of significantly positive or negative temporal correlation between spatial time series data at a great distance.

4.2.2. Application Domains

Discovering various patterns of spatiotemporal coupling and tele-coupling is important in applications related to ecology, environmental science, public safety, and climate science. For example, identifying spatiotemporal cascade patterns from crime event datasets can help police department to understand crime generators in a city, and thus take effective measures to reduce crime events [116].

4.2.3. Statistical Foundation

The underlying statistic for spatiotemporal coupling patterns is the spatiotemporal cross K function [117], which extends spatiotemporal Ripley's K function (Section 3.2) to the case of multiple variables.

4.2.4. Common Approaches

Mixed Drove Spatiotemporal Co-Occurrence Patterns represent subsets of two or more different object-types whose instances are often located in spatial and temporal proximity. Discovering MDCOPs is potentially useful in identifying tactics in battlefields and games, understanding predator-prey interactions, and in transportation (road and network) planning [118,119]. However, mining MDCOPs is computationally very expensive because the interest measures are computationally complex, datasets are larger due to the archival history, and the set of candidate patterns is exponential in the number of object-types. Recent work has produced a monotonic composite interest measure for discovering MDCOPs and novel MDCOP mining algorithms are presented in [48,120]. A filter-and-refine approach has also been proposed to identify spatiotemporal co-occurrence on extended spatial objects [49].

A *spatiotemporal sequential pattern* is a sequence of spatiotemporal event types in the form of $f_1 \rightarrow f_2 \rightarrow \dots \rightarrow f_k$. It represents a "chain reaction" from event type f_1 to event type f_2 and then to event type f_3 until it reaches event type f_k . A spatiotemporal sequential pattern differs from a colocation pattern in that it has a total order of event types. Such patterns are important in applications such as epidemiology where some disease transmission may follow paths between several species through spatial contacts. Mining spatiotemporal sequential patterns is challenging due to the lack of statistically meaningful measures as well as high computation cost. A measure of *sequence index*, which can be interpreted by K-function statistics, was proposed in [52,53], together with computationally efficient

algorithms. Other works have investigated spatiotemporal sequential patterns from data other than spatiotemporal events, such as moving object trajectories [121–123].

Cascading spatio-temporal patterns: Partially ordered subsets of event-types whose instances are located together and occur in stages are called cascading spatio-temporal patterns (CSTP). In the domain of public safety, events such as bar closings and football games are considered generators of crime. Preliminary analysis revealed that football games and bar closing events do indeed generate CSTPs. CSTP discovery can play an important role in disaster planning, climate change science [124,125] (e.g., understanding the effects of climate change and global warming) and public health (e.g., tracking the emergence, spread and re-emergence of multiple infectious diseases [126]). A statistically meaningful metric was proposed to quantify interestingness and computational pruning strategies were proposed to make the pattern discovery process more computationally efficient [50,51].

Spatial time series and tele-connection: Given a collection of spatial time series at different locations, teleconnection discovery aims to identify pairs of spatial time series whose correlation is above a given threshold. Tele-connection patterns are important in understanding oscillations in climate science. Computational challenges arise from the length of the time series and the large number of candidate pairs and the length of time series. An efficient index structure, called a cone-tree, as well as a filter and refine approach [46,127] have been proposed which utilize spatial autocorrelation of nearby spatial time series to filter out redundant pair-wise correlation computation. Another challenge is spurious “high correlation” pairs of locations that happen by chance. Recently, statistical significant tests have been proposed to identify statistically significant tele-connection patterns called dipoles from climate data [47]. The approach uses a “wild bootstrap” to capture the spatio-temporal dependencies, and takes account of the spatial autocorrelation, the seasonality and the trend in the time series over a period of time.

4.3. Spatiotemporal Prediction

4.3.1. What is Spatiotemporal Prediction?

Given spatiotemporal data items, with a set of explanatory variables (also called explanatory attributes or features) and a dependent variable (also called target variables), the *spatiotemporal prediction* problem aims to learn a model that can predict the dependent variable from the explanatory variables. When the dependent variable is discrete, the problem is called *spatiotemporal classification*. When the dependent variable is continuous, the problem is *spatiotemporal regression*. One example of spatiotemporal classification problem is remote sensing image classification over temporal snapshots [128], where the explanatory variables consists of various spectral bands or channels (e.g., blue, green, red, infra-red, thermal, *etc.*) and the dependent variable is a thematic class such as forest, urban, water, and agriculture. Examples of spatiotemporal regression include yearly crop yield prediction [129], and daily temperature prediction at different locations.

4.3.2. Application Domains

Spatiotemporal prediction has broad applications such as land cover classification on remote sensing images [130], future trends projection in global or regional climate variables [131], and real estate price modeling [132].

4.3.3. Statistical Foundation

The statistical foundation of spatiotemporal prediction techniques includes classical statistics augmented to account for lagged (spatially and temporally) variables [133], as well as spatiotemporal statistics including spatial and temporal autocorrelation, spatial heterogeneity and temporal non-stationarity, as well as the multi-scale effect (introduced in the Section 3).

4.3.4. Common Approaches

Spatiotemporal Autoregressive Regression (STAR): In the spatial autoregression model, the spatial dependencies of the error term, or, the dependent variable, are directly modeled in the regression equation [134]. If the dependent values y_i are related to each other, then the regression equation can be modified as $y = \rho W y + X\beta + \epsilon$, where W is the neighborhood relationship contiguity matrix and ρ is a parameter that reflects the strength of the spatial dependencies between the elements of the dependent variable via the logistic function for binary dependent variables. SpatioTemporal Autoregressive Regression (STAR) extends SAR by further explicitly modeling the temporal and spatiotemporal dependency across variables at different locations. More details can be found in [68].

Spatiotemporal Kriging: Kriging [68] is a Geostatistic technique to make predictions at locations where observations are unknown, based on locations where observations are known. In other words, Kriging is a spatial “interpolation” model. Spatial dependency is captured by the spatial covariance matrix, which can be estimated through spatial variograms. Spatiotemporal Kriging [82] generalizes spatial kriging with a spatiotemporal covariance matrix and variograms. It can be used to make predictions from incomplete and noise spatiotemporal data.

Hierarchical Dynamic Spatiotemporal Models: Hierarchical dynamic spatiotemporal models (DSMs) [82], as the name suggests, aim to model spatiotemporal processes dynamically with a Bayesian hierarchical framework. On the top is a data model, which represents the conditional dependency of (actual or potential) observations on the underlying hidden process with latent variables. In the middle is a process model, which captures the spatiotemporal dependency with the process model. On the bottom is a parameter model, which captures the prior distributions of model parameters. DSMs have been widely used in climate science and environment science, e.g., for simulating population growth or atmospheric and oceanic processes. For model inference, Kalman filter can be used under the assumption of linear and Gaussian models.

4.4. Spatiotemporal Partitioning and Summarization

4.4.1. What is Spatiotemporal Partitioning and Summarization?

Spatiotemporal partitioning, or *Spatiotemporal clustering* is the process of grouping similar spatiotemporal data items, and thus partitioning the underlying space and time [27]. It is important in many societal applications. For example, partitioning and summarizing crime data, which is spatial and temporal in nature, helps law enforcement agencies find trends of crimes and effectively deploy their police resources. It is important to note that spatiotemporal partitioning or clustering is closely related to, but not the same as spatiotemporal hotspot detection. Hotspots can be considered as special clusters such that events or activities inside a cluster have much higher intensity than outside.

Spatiotemporal summarization aims to provide a compact representation of spatiotemporal data. For example, traffic accident events on a road network can be summarized into several main routes that cover most of the accidents. Spatiotemporal summarization is often done after or together with spatiotemporal partitioning so that objects in each partition can be summarized by aggregated statistics or representative objects.

4.4.2. Application Domains

Spatiotemporal partitioning and summarization are important in many societal applications such as public safety, public health, and environmental science. For example, partitioning and summarizing crime data, which is spatial and temporal in nature, helps law enforcement agencies find trends of crimes and effectively deploy their police resources [135].

4.4.3. Statistical Foundation

Relevant statistics for spatiotemporal partitioning and summarization include spatiotemporal point density estimation [38] (e.g., Kernel density function), and temporal correlation for spatial time series, *etc.*

4.4.4. Common Approaches

Spatio-Temporal Event Partitioning: Some classic clustering algorithms for two dimensional space can be easily generalized to spatiotemporal scenarios [136]. These techniques can be classified as global partitioning, density-based, hierarchical, and graph based. Global partitioning groups spatial objects to maximize within-group similarity. Examples are K-means, K-Medoids, EM algorithm, CLIQUE [137], BIRCH, and CLARANS [138]. Density based approaches first identify “dense” points and connect them to form contiguous clusters or partitions. Examples include ST-DBSCAN [139], a spatiotemporal extension of its spatial version called DBSCAN [140,141], as well as ST-GRID [142], which splits the space and time into 3D cells and merges dense cells together into clusters. The hierarchical approach, e.g., agglomerative [40], dendrogram [40], and BIRCH [143], partitions or groups spatiotemporal data at different hierarchical levels. Graph based approaches such as Chameleon [40,144] first represent data

items as a sparse k nearest neighbor graph, then partition the graph into segments, and hierarchically merge graph segments according to the similarity between original segments and merged segments.

Spatial Time-Series Partitioning: Spatial time series partitioning aims to divide the space into regions such that the correlation or similarity between time series within the same regions is maximized. Global partitioning methods such as K Means, K Medoids, and EM can be applied. So can the hierarchical approach. However, due to the high dimensionality of spatial time series, density-based approaches and graph-based approaches are often not effective. When computing similarity between spatial time series, a filter-and-refine approach [46] can be used to avoid redundant computation.

Trajectory Data Partitioning: Trajectory data partitioning aims to partition trajectories into groups according to their similarity. Algorithms are of two types, *i.e.*, density-based and frequency-based. The *density based approaches* [145,146] first break trajectories into small segments and apply density-based clustering algorithms similar to DB-SCAN [140] to connect dense areas of segments. The *frequency based approach* [147] uses association rule mining [63] algorithms to identify subsections of trajectories which have high frequencies (also called high “support”).

Spatiotemporal Summarization: Data summarization aims to find compact representation of a data set [148]. It is important for data compression as well as for making pattern analysis more convenient. As shown in Table 5, data summarization can be conducted on classical data, spatial data, as well as spatiotemporal data. For spatial time series data, summarization can be done by removing spatial and temporal redundancy due to the effect of autocorrelation. A family of such algorithms has been used to summarize traffic data streams [149]. Similarly, the centroids from K Means can also be used to summarize spatial time series. For trajectory data, especially spatial network trajectories, summarization is more challenging due to the huge cost of similarity computation. A recent approach summarizes network trajectories into k primary corridors [150,151]. The work proposes efficient algorithms to reduce the huge cost for network trajectory distance computation.

Table 5. Summarization Framework for Various Data Types.

Data Types	Partition Definition	Summarization
classical data	partition of rows of records	aggregate statistics: sum, count, mean, <i>etc.</i>
spatial data	partition of Euclidean space	representatives: centroids, medoids, <i>etc.</i>
	partition of spatial network	representatives: K main routes, <i>etc.</i>
spatio-temporal data	partition of trajectories on a spatial or spatio-temporal network	representatives: K primary corridors, <i>etc.</i>

4.5. Spatiotemporal Hotspots

4.5.1. What are Spatiotemporal Hotspots?

Given a set of spatial objects (e.g., activity locations) in a study area, *spatiotemporal hotspots* are regions together certain time intervals where the number of objects is anomalously or unexpectedly high

within the time intervals. Spatiotemporal hotspots are a special kind of clustered pattern whose inside has significantly higher intensity than outside.

4.5.2. Application Domains

Application domains for spatiotemporal hotspot detection range from public health to criminology. For example, in epidemiology finding disease hotspots allows officials to detect an epidemic and allocate resources to limit its spread [152].

4.5.3. Statistical Foundation

Spatiotemporal scan statistics [76,152] are used to detect statistically significant hotspots from spatiotemporal datasets. It uses a cylinder to scan the space-time for candidate hotspots and perform hypothesis testing. The null hypothesis states that the activity points are distributed randomly according to a homogeneous (*i.e.*, same intensity) Poisson process over the geographical space. The alternative hypothesis states that the inside of the cylinder has higher intensity of activities than outside. A test statistic called the log likelihood ratio is computed for each candidate hotspot (or cylinder) and the candidate with the highest likelihood ratio can be evaluated using a significance value (*i.e.*, p -value).

4.5.4. Common Approaches

Clustering based approaches: Clustering methods can be used to identify candidate areas for a further evaluation of spatiotemporal hotspots. These methods include global partitioning, density based clustering and hierarchical clustering (see Section 4.4). These methods can be used as a preprocessing step to generate candidate hotspot areas, and statistical tools may be used to test statistical significance. CrimeStat, a software package for spatial analysis of crime locations, incorporates several clustering methods to determine the crime hotspots in a study area. CrimeStat package has k-means tool, nearest neighbor hierarchical (NNH) clustering [153], Risk Adjusted NNH (RANNH) tool, STAC Hot Spot Area tool [135], and a Local Indicator of Spatial Association (LISA) tool [72] that are used to evaluate potential hotspot areas. Although many of the clustering methods mentioned above are generally designed for two dimensional Euclidean space and mostly used for pure spatial data, they can be used to identify spatiotemporal candidate hotspots by considering the temporal part of the data as a third dimension. For example, DBSCAN will cluster both spatial and spatiotemporal data using the density of the data as its measure.

Spatiotemporal Scan Statistics based approaches: Spatiotemporal hotspot detection can be seen as a special case of pure spatial hotspot detection by the addition of time as a third dimension. Two types of spatiotemporal hotspots that are of particular importance: “*persistent*” *spatiotemporal hotspot* and “*emerging*” *spatiotemporal hotspot*. A “persistent” spatiotemporal hotspot is defined as a region where the rate of increase in observations is constantly high over time. Thus, a persistent hotspot detection assumes that the risk of an hotspot (*i.e.*, outbreak) is constant over time and it searches over space and time for an hotspot by simply totaling the number of observations in each time interval. An example tool for persistent spatiotemporal hotspot detection is SaTScan, which uses a cylindrical window in three dimensions (time is the third dimension) instead of the circular window in two dimensions [152]

used for detecting spatial hotspots. An “emerging” spatiotemporal hotspot is a region where the rate of observations is monotonically increasing over time [154,155]. This kind of spatiotemporal hotspot occurs when an outbreak emerges causing a sudden increase in the number observations. Such phenomena can be observed in epidemiology where at the start of an outbreak the number of disease cases suddenly increases. Tools for the detection of emerging spatiotemporal hotspots use spatial scan statistics with the change in expectation over time [156].

4.6. Spatiotemporal Change

4.6.1. What are Spatiotemporal Changes and Change Footprints

Although the single term “change” is used to name the spatiotemporal change footprint patterns in different applications, the underlying phenomena may differ significantly. This section briefly summarizes the main ways a change may be defined in spatiotemporal data [29]:

Change in Statistical Parameter: In this case, the data is assumed to follow a certain distribution and the change is defined as a shift in this statistical distribution. For example, in statistical quality control, a change in the mean or variance of the sensor readings is used to detect a fault.

Change in Actual Value: Here, change is modeled as the difference between a data value and its spatial or temporal neighborhood. For example, in a one-dimensional continuous function, the magnitude of change can be characterized by the derivative function, while on a two-dimensional surface, it can be characterized by the gradient magnitude.

Change in Models Fitted to Data: This type of change is identified when a number of function models are fitted to the data and one or more of the models exhibits a change (e.g., a discontinuity between consecutive linear functions) [157].

4.6.2. Common Approaches

This section follows the taxonomy of spatiotemporal change footprint patterns as proposed in [29]. In this taxonomy, spatiotemporal change footprints are classified along two dimensions: temporal and spatial. Temporal footprints are classified into four categories: single snapshot, set of snapshots, point in a long series, and interval in a long series. Single snapshot refers to a purely spatial change that does not have a temporal context. A set of snapshots indicates a change between two or more snapshots of the same spatial field, e.g., satellite images of the same region.

Spatial footprints can be classified as raster footprints or vector footprints. Vector footprints are further classified into four categories: point(s), line(s), polygon(s), and network footprint patterns. Raster footprints are classified based on the scale of the pattern, namely, local, focal, or zonal patterns. This classification describes the scale of the change operation of a given phenomenon in the spatial raster field [158]. Local patterns are patterns in which change at a given location depends only on attributes at this location. Focal patterns are patterns in which change in a location depends on attributes in that location and its assumed neighborhood. Zonal patterns define change using an aggregation of location values in a region.

Spatiotemporal Change Patterns with Raster-Based Spatial Footprint: This includes patterns of *spatial changes between snapshots*. In remote sensing, detecting changes between satellite images can help identify land cover change due to human activity, natural disasters, or climate change [159–161]. Given two geographically aligned raster images, this problem aims to find a collection of pixels that have significant changes between the two images [162]. This pattern is classified as a local change between snapshots since the change at a given pixel is assumed to be independent of changes at other pixels. Alternative definitions have assumed that a change at a pixel also depends on its neighborhoods [163]. For example, the pixel values in each block may be assumed to follow a Gaussian distribution [164]. We refer to this type of change footprint pattern as a focal spatial change between snapshots. Researchers in remote sensing and image processing have also tried to apply image change detection to objects instead of pixels [165–167], yielding zonal spatial change patterns between snapshots.

A well-known technique for detecting a local change footprint is simple differencing. The technique starts by calculating the differences between the corresponding pixels intensities in the two images. A change at a pixel is flagged if the difference at the pixel exceeds a certain threshold. Alternative approaches have also been proposed to discover focal change footprints between images. For example, the block-based density ratio test detects change based on a group of pixels, known as a block [168,169]. Object-based approaches in remote sensing [167,170,171] employ image segmentation techniques to partition temporal snapshots of images into homogeneous objects [172] and then classify object pairs in the two temporal snapshots of images into no change or change classes.

Spatiotemporal Change Patterns with Vector-Based Spatial Footprint: This includes the *Spatiotemporal Volume Change Footprint* pattern. This pattern represents a change process occurring in a spatial region (a polygon) during a time interval. For example, an outbreak event of a disease can be defined as an increase in disease reports in a certain region during a certain time window up to the current time. Change patterns known to have an spatiotemporal volume footprint include the spatiotemporal scan statistics [173,174], a generalization of the spatial scan statistic, and emerging spatiotemporal clusters defined by [156].

5. Spatial and Spatiotemporal Analysis Tools

This section lists currently existing spatial and spatiotemporal analysis tools, including geographic information system (GIS) softwares, spatial and spatiotemporal statistical tools, spatial database management systems, as well as spatial big data platforms.

GIS Softwares: ArcGIS [175] is the currently most widely used commercial GIS software for working with maps and geographic information. It has an extension named Tracking Analyst to support visualization and analysis for spatiotemporal data. QGIS [176] (previously Quantum GIS) is a very popular open source GIS software.

Spatial Statistical Tools: R provides many packages for spatial and spatiotemporal statistical analysis [177], such as *spatstat* for point pattern analysis, *gstat* and *geoR* for Geostatistics, *spdep* for areal data analysis. Matlab also provides Mapping Toolbox [178] and other spatial statistical toolboxes. SAS recently provides support on spatial statistics [179] such as KRIGE2D Procedure for Kriging, SIM2D

Procedure for Gaussian random field, SPP Procedure for spatial point pattern, and VARIOGRAM Procedure for variograms.

Spatial Database Management Systems: Many commercial database provides extensions to support spatial data, such as Oracle Spatial [180], and DB2 Spatial Extender [181]. PostGIS [182] is a widely used open source spatial database management systems, which is an extension to Postgres, an object-relational DBMS.

Spatial Big Data Platform: The upcoming spatial big data from vehicle GPS trajectories, cellphone location data, as well as remote sensing imagery exceeds the capabilities of traditional spatial DBMS, and requires new platforms to support scalable spatial analysis. Current spatial big data platforms include ESRI GIS on Hadoop [183,184], Hadoop GIS [185], and Spatial Hadoop [186].

6. Research Trend and Future Research Needs

Most current research in spatiotemporal data mining uses Euclidean space, which often assumes isotropic property and symmetric neighborhoods. However, in many real world applications, the underlying space is network space, such as river networks and road networks [187–189]. One of the main challenges in spatial and spatiotemporal network data mining is to account for the network structure in the dataset. For example, in anomaly detection, spatial techniques do not consider the spatial network structure of the dataset, that is, they may not be able to model graph properties such as one-ways, connectivity, left-turns, *etc.* The network structure often violates the isotropic property and symmetry of neighborhoods, and instead, requires asymmetric neighborhood and directionality of neighborhood relationship (e.g., network flow direction).

Recently, some cutting edge research has been conducted in the spatial network statistics and data mining [80]. For example, several spatial network statistical methods have been developed, e.g., network K function and network spatial autocorrelation. Several spatial analysis methods have also been generalized to the network space, such as network point cluster analysis and clumping method, network point density estimation, network spatial interpolation (Kriging), as well as network Huff model. Due to the nature of spatial network space as distinct from Euclidean space, these statistics and analysis often rely on advanced spatial network computational techniques [80].

We believe more spatiotemporal data mining research is still needed in the network space. First, though several spatial statistics and data mining techniques have been generalized to the network space, few spatiotemporal network statistics and data mining have been developed, and the vast majority of research is still in the Euclidean space. Future research is needed to develop more spatial network statistics, such as spatial network scan statistics, spatial network random field model, as well as spatiotemporal autoregressive models for networks. Furthermore, phenomena observed on spatiotemporal networks need to be interpreted in an appropriate frame of reference to prevent a mismatch between the nature of the observed phenomena and the mining algorithm. For instance, moving objects on a spatiotemporal network need to be studied from a traveler's perspective, *i.e.*, the Lagrangian frame of reference [190–192] instead of a snapshot view. This is because a traveler moving along a chosen path in a spatiotemporal network would experience a road-segment (and its properties such as fuel efficiency, travel-time *etc.*) for the time at which he/she arrives at that segment, which

may be distinct from the original departure-time at the start of the journey. These unique requirements (non-isotropy and Lagrangian reference frame) call for novel spatiotemporal statistical foundations [187] as well as new computational approaches for spatiotemporal network data mining.

Another future research need is to develop spatiotemporal graph big data platforms, motivated by the upcoming rich spatiotemporal network data collected from vehicles. Modern vehicles have rich instrumentation to measure hundreds of attributes at high frequency and are generating big data (Exabyte [193]). This vehicle measurement big data (VMBD) consist of a collection of trips on a transportation graph such as a road map annotated with several measurements of engine sub-systems. Collecting and analyzing VMBD during real-world driving conditions can aid in understanding the underlying factors which govern real world fuel inefficiencies or high greenhouse gas (GHG) emissions [194]. Current relevant big data platforms for spatial and spatiotemporal data mining include ESRI GIS Tools for Hadoop [183,184], Hadoop GIS [185], *etc.* These provide distributed systems for geometric-data (e.g., lines, points and polygons) including geometric indexing and partitioning methods such as R-tree, R+-tree, or Quad tree. Recently, SpatialHadoop has been developed [186]. SpatialHadoop embeds geometric notions in language, visualization, storage, MapReduce, and operations layers. However, spatio-temporal graphs (STGs) violate the core assumptions of current spatial big data platforms that the geometric concepts are adequate for conveniently representing STG analytics operations and for partition data for load-balancing. STGs also violate core assumptions underlying graph analytics software (e.g., Giraph [195], GraphLab [196] and Pregel [197]) that traditional location-unaware graphs are adequate for conveniently representing STG analytics operations and for partition data for load-balancing. Therefore, novel spatiotemporal graph big data platforms are needed. Several challenges should be addressed, e.g., spatiotemporal graph big data requires novel distributed file systems (DFS) to partition the graph, and a novel programming model is still needed to support abstract data types and fundamental STG operations, *etc.*

7. Summary

This paper provides an over view of current research in the field of spatiotemporal data mining from a computational perspective. Spatiotemporal data mining has broad application domains including ecology and environmental management, public safety, transportation, earth science, epidemiology, and climatology. However, the complexity of spatiotemporal data and intrinsic relationships limits the usefulness of conventional data science techniques for extracting spatiotemporal patterns. We provide a taxonomy of different spatiotemporal data types and underlying spatiotemporal statistics. We also review common spatiotemporal data mining techniques organized by major output pattern families: spatiotemporal outlier, spatiotemporal coupling and tele-coupling, spatiotemporal prediction, spatiotemporal partitioning and summarization, spatiotemporal hotspots, and change detection. Popular software tools for spatial and spatiotemporal data analysis are also listed. Finally, we discuss the cutting edge research areas and future research needs.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1029711, IIS-1320580, 0940818 and IIS-1218168, the USDOD under Grant No. HM1582-08-1-0017 and HM0210-13-1-0005, and the University of Minnesota under OVPR U-Spatial. We would like to thank Kim Koffolt for the helpful comments in improving the readability of the paper. We would also like to thank Pradeep Mohan for helpful advice.

Author Contributions

Shashi Shekhar provided materials from previous surveys or tutorials for the early draft, added new references, and provided overall guidance. Zhe Jiang contributed to the organization of the paper, the writing of Sections 1–3, 4.1–4.3, and 5, and coordinating with other authors. He also took charge of correspondence during the paper submission, handling the revisions and re-submission of revised manuscripts. Reem Y. Ali contributed to Section 4.6 and Section 6. Emre Eftelioglu contributed to Section 4.5. Xun Tang contributed to Section 4.4. Venkata M. V. Gunturi contributed to Section 6. Xun Zhou contributed to Section 4.6.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Stolorz, P.; Nakamura, H.; Mesrobian, E.; Muntz, R.; Shek, E.; Santos, J.; Yi, J.; Ng, K.; Chien, S.; Mechoso, R.; *et al.* *Fast Spatio-Temporal Data Mining of Large Geophysical Datasets*; AAAI Press: Palo Alto, CA, USA, 1995.
2. Guting, R. An introduction to spatial database systems. *VLDB J.* **1994**, *3*, 357–399.
3. Shekhar, S.; Chawla, S. *Spatial Databases: A Tour*; Prentice Hall: Upper Saddle River, NJ, USA, 2003.
4. Shekhar, S.; Chawla, S.; Ravada, S.; Fetterer, A.; Liu, X.; Lu, C.T. Spatial databases—Accomplishments and research needs. *Trans. Knowl. Data Eng.* **1999**, *11*, 45–55.
5. Worboys, M. *GIS: A Computing Perspective*; Taylor and Francis: London, UK, 1995.
6. Krugman, P. *Development, Geography, and Economic Theory*; MIT Press: Cambridge, MA, USA, 1995.
7. Albert, P.; McShane, L. A generalized estimating equations approach for spatially correlated binary data: Applications to the analysis of neuroimaging data. *Biometrics* **1995**, *51*, 627–638.
8. Shekhar, S.; Yang, T.; Hancock, P. An intelligent vehicle highway information management system. *Comput.—Aided Civil Infrastruct. Eng.* **1993**, *8*, 175–198.
9. Eck, J. E.; Chainey, S.; Cameron, J. G.; Leitner, M.; Wilson, R. E. Mapping Crime: Understanding Hot Spots. Available online: <http://www.ncjrs.gov/pdffiles1/nij/209393.pdf> (accessed on 10 May 2015).
10. Issaks, E.H.; Svivastava, R.M. *Applied Geostatistics*; Oxford University Press: Oxford, UK, 1989.

11. Haining, R.J. *Spatial Data Analysis in the Social and Environmental Sciences*; Cambridge University Press: Cambridge, UK, 1989.
12. Roddick, J.F.; Spiliopoulou, M. A bibliography of temporal, spatial and spatio-temporal data mining research. *SIGKDD Explor.* **1999**, *1*, 34–38 .
13. Scally, R. *GIS for Environmental Management*; ESRI Press: Redlands, CA, USA, 2006.
14. Leipnik, M.R.; Albert, D.P. *GIS in Law Enforcement: Implementation Issues and Case Studies*; CRC Press: Sacramento, CA, USA, 2002.
15. Lang, L. *Transportation GIS*; ESRI Press: Redlands, CA, USA, 1999.
16. Elliott, P.; Wakefield, J.; Best, N.; Briggs, D. *Spatial Epidemiology: Methods and Applications*; Oxford University Press: Oxford, UK, 2000.
17. Hohn, M.; A.E. Liebhold, L.G. A Geostatistical model for forecasting the spatial dynamics of defoliation caused by the Gypsy Moth, *Lymantria dispar* (Lepidoptera:Lymantriidae). *Environ. Entomol.* **1993**, *22*, 1066–1075.
18. Yasui, Y.; Lele, S. A regression method for spatial disease rates: An estimating function approach. *J. Am. Stat. Assoc.* **1997**, *94*, 21–32.
19. Ruß, G.; Brenning, A. Data mining in precision agriculture: Management of spatial information. In *Computational Intelligence for Knowledge-Based Systems Design*; Springer: Berlin, Germany, 2010; pp. 350–359.
20. Gubbi, J.; Buyya, R.; Marusic, S.; Palaniswami, M. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Gener. Comput. Syst.* **2013**, *29*, 1645–1660.
21. Marcus, G.; Davis, E. Eight (no, nine!) problems with big data. *N. Y. Times* **2014**, *6*, 2014.
22. Caldwell, P.M.; Bretherton, C.S.; Zelinka, M.D.; Klein, S.A.; Santer, B.D.; Sanderson, B.M. Statistical significance of climate sensitivity predictors obtained by data mining. *Geophys. Res. Lett.* **2014**, *41*, 1803–1808.
23. Shekhar, S.; Zhang, P.; Huang, Y.; Vatsavai, R.R. Trends in spatial data mining. In *Data Mining: Next Generation Challenges and Future Directions*; AAAI Press: Palo Alto, CA, USA, 2003; pp. 357–380.
24. Koperski, K.; Adhikary, J.; Han, J. Spatial data mining: Progress and challenges survey paper. In Proceedings of the ACM SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery, Montreal, QC, Canada, 4–6 June 1996.
25. Ester, M.; Kriegel, H.P.; Sander, J. Spatial Data Mining: A Database Approach. In *Advances in Spatial Databases*, Proceedings of the 5th International Symposium (SSD '97), Berlin, Germany, 15–18 July 1997; Springer: Berlin, Germany, 1997; pp. 47–66.
26. Miller, H.J.; Han, J. *Geographic Data Mining and Knowledge Discovery*; CRC Press: Sacramento, CA, USA, 2009.
27. Kisilevich, S.; Mansmann, F.; Nanni, M.; Rinzivillo, S. *Spatio-Temporal Clustering*; Springer: Berlin, Germany, 2010.
28. Aggarwal, C.C. *Outlier Analysis*; Springer: Berlin, Germany, 2013.
29. Zhou, X.; Shekhar, S.; Ali, R.Y. Spatiotemporal change footprint pattern discovery: An inter-disciplinary survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2014**, *4*, 1–23.

30. Cheng, T.; Haworth, J.; Anbaroglu, B.; Tanaksaranond, G.; Wang, J. Spatiotemporal data mining. In *Handbook of Regional Science*; Springer: Heidelberg, Germany, 2014; pp. 1173–1193.
31. Shekhar, S.; Evans, M.R.; Kang, J.M.; Mohan, P. Identifying patterns in spatial information: A survey of methods. *Wiley Interdiscip. Rev.: Data Min. Knowl. Discov.* **2011**, *1*, 193–214.
32. Worboys, M.; Duckham, M. *GIS: A Computing Perspective*, 2nd ed.; CRC Press: Sacramento, CA, USA, 2004.
33. Li, Z.; Chen, J.; Baltsavias, E. *Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences: 2008 ISPRS Congress Book*; CRC Press: Sacramento, CA, USA, 2008.
34. Yuan, M. Temporal GIS and spatio-temporal modeling. In Proceedings of the Third International Conference Workshop on Integrating GIS and Environment Modeling, Santa Fe, NM, USA, 21–26 January 1996.
35. Allen, J.F. Towards a general theory of action and time. *Artif. Intell.* **1984**, *23*, 123–154.
36. George, B.; Kim, S.; Shekhar, S. Spatio-temporal Network Databases and Routing Algorithms: A Summary of Results. In Proceedings of the 10th International Symposium on Spatial and Temporal Databases (SSTD'07), Boston, MA, USA, 16–18 July 2007.
37. George, B.; Shekhar, S. Time Aggregated Graphs: A model for spatio-temporal network. In Proceedings of the Workshops (CoMoGIS) at the 25th International Conference on Conceptual Modeling (ER2006), Tucson, AZ, USA, 6–9 November 2006.
38. Gelfand, A.E.; Diggle, P.; Guttorp, P.; Fuentes, M. *Handbook of Spatial Statistics*; CRC Press: Sacramento, CA, USA, 2010.
39. Campelo, C.E.; Bennett, B. *Representing and Reasoning about Changing Spatial Extensions of Geographic Features*; Springer: Berlin, Germany, 2013.
40. Tan, P.N.; Steinbach, M.; Kumar, V. *Introduction to Data Mining*; Pearson Addison Wesley: Boston, MA, USA, 2006.
41. Bolstad, P. *GIS Fundamentals: A First Text on GIS*; Eider Press: Saint Paul, MN, USA, 2002.
42. Ganguly, A.R.; Steinhäuser, K. Data Mining for Climate Change and Impacts. In Proceedings of the 2008 IEEE International Conference on Data Mining Workshops (ICDMW '08), Pisa, Italy, 15–19 December 2008; pp. 385–394.
43. Erwig, M.; Schneider, M.; Hagen, F. Spatio-temporal predicates. *IEEE Trans. Knowl. Data Eng.* **2002**, *14*, 881–901.
44. Chen, J.; Wang, R.; Liu, L.; Song, J. Clustering of trajectories based on Hausdorff distance. In Proceedings of the 2011 IEEE International Conference on Electronics, Communications and Control (ICECC), Ningbo, China, 9–11 September 2011; pp. 1940–1944.
45. Zhang, Z.; Huang, K.; Tan, T. Comparison of similarity measures for trajectory clustering in outdoor surveillance scenes. In Proceedings of the 18th IEEE International Conference on Pattern Recognition (ICPR 2006), Hong Kong, China, 20–24 August 2006; Volume 3, pp. 1135–1138.
46. Zhang, P.; Huang, Y.; Shekhar, S.; Kumar, V. Correlation analysis of spatial time series datasets: A filter-and-refine approach. In *Advances in Knowledge Discovery and Data Mining*; Springer: Berlin, Germany, 2003; pp. 532–544.

47. Kawale, J.; Chatterjee, S.; Ormsby, D.; Steinhaeuser, K.; Liess, S.; Kumar, V. Testing the significance of spatio-temporal teleconnection patterns. In Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, China, 12–16 August 2012; pp. 642–650.
48. Celik, M.; Shekhar, S.; Rogers, J.P.; Shine, J.A.; Yoo, J.S. Mixed-drove spatio-temporal co-occurrence pattern mining: A summary of results. In Proceedings of the Sixth International Conference on Data Mining; Washington, DC, USA, 18–22 December 2006.
49. Pillai, K.G.; Angryk, R.A.; Aydin, B. *A Filter-and-Refine Approach to Mine Spatiotemporal Co-Occurrences*; SIGSPATIAL/GIS: Orlando, USA, 2013; pp. 104–113.
50. Mohan, P.; Shekhar, S.; Shine, J.A.; Rogers, J.P. Cascading spatio-temporal pattern discovery. *IEEE Trans. Knowl. Data Eng.* **2012**, *24*, 1977–1992.
51. Mohan, P.; Shekhar, S.; Shine, J.A.; Rogers, J.P. *Cascading Spatio-Temporal Pattern Discovery: A Summary of Results*; SDM: Columbus, USA, 2010.
52. Huang, Y.; Zhang, L.; Zhang, P. A framework for mining sequential patterns from spatio-temporal event data sets. *IEEE Trans. Knowl. Data Eng.* **2008**, *20*, 433–448.
53. Huang, Y.; Zhang, L.; Zhang, P. *Finding Sequential Patterns from a Massive Number of Spatio-Temporal Events*; SDM: Bethesda, USA, 2006.
54. Mennis, J.; Viger, R.; Tomlin, C.D. Cubic map algebra functions for spatio-temporal analysis. *Cartogr. Geogr. Inf. Sci.* **2005**, *32*, 17–32.
55. Brown, D.G.; Riolo, R.; Robinson, D.T.; North, M.; Rand, W. Spatial process and data models: Toward integration of agent-based models and GIS. *J. Geogr. Syst.* **2005**, *7*, 25–47.
56. Habiba, H.; Tantipathananandh, C.; Berger-Wolf, T. *Betweenness Centrality Measure in Dynamic Networks*; Department of Computer Science, University of Illinois at Chicago: Chicago, IL, USA, 2007.
57. Kang, J.M.; Shekhar, S.; Wennen, C.; Novak, P. Discovering flow anomalies: A SWEET approach. In Proceedings of the 8th IEEE International Conference on Data Mining (ICDM '08), Pisa, Italy, 15–19 December 2008.
58. Kang, J.M.; Shekhar, S.; Henjum, M.; Novak, P.; Arnold, W. Discovering teleconnected flow anomalies: A relationship analysis of spatio-temporal Dynamic (RAD) neighborhoods. In Proceedings of 11th International Symposium (SSTD 2009), Aalborg, Denmark, 8–10 July 2009.
59. Ahuja, R.K.; Magnanti, T.L.; Orlin, J.B. *Network Flows: Theory, Algorithms, and Applications*; Prentice Hall: Upper Saddle River, NJ, USA, 1988.
60. Quinlan, J. *C4.5: Programs for Machine Learning*; Morgan Kaufmann Publishers: Burlington, MA, USA, 1993.
61. Varnett, V.; Lewis, T. *Outliers in Statistical Data*; John Wiley: Hoboken, NJ, USA, 1994.
62. Agarwal, R.; Imielinski, T.; Swami, A. Mining association rules between sets of items in large databases. In Proceedings of the ACM SIGMOD Conference on Management of Data, Washington, DC, USA, 26–28 May 1993.
63. Agrawal, R.; Srikant, R. Fast algorithms for mining association rules. In Proceedings of the 20th International Conference on Very Large Data Bases, Santiago de Chile, Chile, 12–15 September 1994.

64. Jain, A.; Dubes, R. *Algorithms for Clustering Data*; Prentice Hall: Upper Saddle River, NJ, USA, 1988.
65. Tobler, W. *Cellular Geography, Philosophy in Geography*; Reidel: Dordrecht, Netherlands, 1979.
66. Banerjee, S.; Carlin, B.; Gelfand, A. *Hierarchical Modeling and Analysis for Spatial Data*; Chapman & Hall: London, UK, 2004.
67. Schabenberger, O.; Gotway, C. *Statistical Methods for Spatial Data Analysis*; Chapman and Hall: London, UK, 2005.
68. Cressie, N.A.C. *Statistics for Spatial Data*; Wiley: New York, NY, USA, 1993.
69. Cressie, N. *Statistics for Spatial Data (Revised Edition)*; Wiley: New York, NY, USA, 1993.
70. Gething, P.W.; Atkinson, P.M.; Noor, A.; Gikandi, P.; Hay, S.I.; Nixon, M.S. A local space-time kriging approach applied to a national outpatient malaria data set. *Comput. Geosci.* **2007**, *33*, 1337–1350.
71. Warrender, C.E.; Augusteijn, M.F. Fusion of image classifications using Bayesian techniques with Markov random fields. *Int. J. Remote Sens.* **1999**, *20*, 1987–2002.
72. Anselin, L. Local indicators of spatial association-LISA. *Geogr. Anal.* **1995**, *27*, 93–155.
73. Openshaw, S. *The Modifiable Areal Unit Problem*; OCLC: Dublin, OH, USA, 1983.
74. Ripley, B.D. Modelling spatial patterns. *J. R. Stat. Soc. Ser. B (Methodol.)* **1977**, *39*, 172–212.
75. Marcon, E.; Puech, F. *Generalizing Ripley's K Function to Inhomogeneous Populations*; Mimeo: New York, NY, USA, 2003.
76. Kulldorff, M. A spatial scan statistic. *Commun. Stat.-Theory Methods* **1997**, *26*, 1481–1496.
77. Chiu, S.N.; Stoyan, D.; Kendall, W.S.; Mecke, J. *Stochastic Geometry and Its Applications*; John Wiley & Sons: New York, NY, USA, 2013.
78. Guyon, X. *Random Fields on a Network: Modeling, Statistics, and Applications*; Springer: Berlin, Germany, 1995.
79. Okabe, A.; Yomono, H.; Kitamura, M. Statistical analysis of the distribution of points on a network. *Geogr. Anal.* **1995**, *27*, 152–175.
80. Okabe, A.; Sugihara, K. *Spatial Analysis along Networks: Statistical and Computational Methods*; John Wiley & Sons: New York, NY, USA, 2012.
81. Okabe, A.; Okunuki, K.; Shiode, S. The SANET Toolbox: New methods for network spatial analysis. *Trans. GIS* **2006**, *10*, 535–550.
82. Cressie, N.; Wikle, C.K. *Statistics for Spatio-Temporal Data*; John Wiley & Sons: New York, NY, USA, 2011.
83. Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and Its Applications: With R Examples*; Springer: Berlin, Germany, 2010.
84. Kyriakidis, P.C.; Journel, A.G. Geostatistical space-time models: A review. *Math. Geol.* **1999**, *31*, 651–684.
85. Cressie, A.C.N. *Statistics for Spatial Data*; Wiley-Interscience: Hoboken, NJ, USA, 1993.
86. Barnett, V.; Lewis, T. *Outliers in Statistical Data*, 3rd ed.; John Wiley: Hoboken, NJ, USA, 1994.
87. Hawkins, D. *Identification of Outliers*; Chapman and Hall: London, UK, 1980.
88. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly Detection: A Survey. *ACM Comput. Surv.* **2009**, *41*, 1–58.

89. Shekhar, S.; Lu, C.; Zhang, P. Graph-based outlier detection: Algorithms and applications (a summary of results). In Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 26–29 August 2001.
90. Wu, W.; Cheng, X.; Ding, M.; Xing, K.; Liu, F.; Deng, P. Localized outlying and boundary data detection in sensor networks. *IEEE Trans. Knowl. Data Eng.* **2007**, *19*, 1145–1157.
91. Sun, P.; Chawla, S. On local spatial outliers. In Proceedings of the Fourth IEEE International Conference on Data Mining (ICDM '04), Brighton, UK, 1–4 November 2004; pp. 209–216.
92. Pei, Y.; Zaiane, O.R.; Gao, Y. An efficient reference-based approach to outlier detection in large Datasets. In Proceedings of the Sixth International Conference on Data Mining (ICDM '06), Hong Kong, China, 18–22 December 2006; pp. 478–487.
93. Shekhar, S.; Lu, C.; Zhang, P. A unified approach to detecting spatial outliers. *GeoInformatica* **2003**, *7*, 139–166.
94. Shekhar, S.; Lu, C.; Zhang, P. Detecting graph-based spatial outliers: Algorithms and applications. In Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 26–29 August 2001.
95. Haslett, J.; Bradley, R.; Craig, P.; Unwin, A.; Wills, G. Dynamic graphics for exploring spatial data with application to locating global and local anomalies. *Am. Stat.* **1991**, *45*, 234–242.
96. Luc, A. Local indicators of spatial association: LISA. *Geogr. Anal.* **1995**, *27*, 93–115.
97. Luc, A. Exploratory spatial data analysis and geographic information systems. In *New Tools for Spatial Analysis*; Wiley: Hoboken, NJ, USA, 1994; pp. 45–54.
98. Chen, F.; Lu, C.T.; Boedihardjo, A.P. *GLS-SOD: A Generalized Local Statistical Approach for Spatial Outlier Detection*; KDD: Washington, DC, USA, 2010.
99. Lu, C.T.; Chen, D.; Kou, Y. Algorithms for spatial outlier detection; In Proceedings of the 3rd International Conference on Data Mining (ICDM '03), Melbourne, FL, USA, 19–22 November 2003; pp. 597–600.
100. Chen, D.; Lu, C.T.; Kou, Y.; Chen, F. On detecting spatial outliers. *GeoInformatica* **2008**, *12*, 455–475.
101. McGuire, M.P.; Janeja, V.P.; Gangopadhyay, A. Mining trajectories of moving dynamic spatio-temporal regions in sensor datasets. *Data Min. Knowl. Discov.* **2014**, *28*, 961–1003.
102. Lu, C.T.; Chen, D.; Kou, Y. Detecting spatial outliers with multiple attributes. In Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence; Washington, DC, USA, 3–5 November 2003.
103. Kou, Y.; Lu, C.T.; Chen, D. *Spatial Weighted Outlier Detection*; SDM: Bethesda, USA, 2006.
104. Liu, X.; Chen, F.; Lu, C.T. On detecting spatial categorical outliers. *GeoInformatica* **2014**, *18*, 501–536.
105. Schubert, E.; Zimek, A.; Kriegel, H.P. Local outlier detection reconsidered: A generalized view on locality with applications to spatial, video, and network outlier detection. *Data Min. Knowl. Discov.* **2014**, *28*, 190–237.
106. Wu, M.; Song, X.; Jermaine, C.; Ranka, S.; Gums, J. *A LRT Framework for Fast Spatial Anomaly Detection*; KDD: Paris, France, 2009.

107. Wu, M.; Jermaine, C.; Ranka, S.; Song, X.; Gums, J. A model-agnostic framework for fast spatial anomaly detection. *TKDD* **2010**, *4*, doi:10.1145/1857947.1857952.
108. Franke, C.; Gertz, M. Detection and exploration of outlier regions in sensor data streams. In Proceedings of the IEEE International Conference on Data Mining Workshops (ICDMW '08), Pisa, Italy, 15–19 December 2008; pp. 375–384.
109. Elfeky, M.G.; Aref, W.G.; Elmagarmid, A.K. STAGGER: Periodicity mining of data streams using expanding sliding windows. In Proceedings of the Sixth International Conference on Data Mining (ICDM '06), Hong Kong, China, 18–22 December 2006; pp. 188–199.
110. Liu, C.; Xiong, H.; Ge, Y.; Geng, W.; Perkins, M. A stochastic model for context-aware anomaly detection in indoor location traces. In Proceedings of the 12th International Conference on Data Mining (ICDM '12), Brussels, Belgium, 10–13 December 2012; pp. 449–458.
111. Bu, Y.; Chen, L.; Fu, A.W.C.; Liu, D. Efficient anomaly monitoring over moving object trajectory streams. In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, 28 June–1 July 2009; pp. 159–168.
112. Li, X.; Han, J.; Kim, S. Motion-alert: Automatic anomaly detection in massive moving objects. In *Intelligence and Security Informatics*; Springer: Berlin, Germany, 2006; pp. 166–177.
113. Ge, Y.; Xiong, H.; Liu, C.; Zhou, Z.H. A taxi driving fraud detection system. In Proceedings of the 2011 IEEE International Conference on Data Mining (ICDM), Vancouver, BC, Canada, 11–14 December 2011; pp. 181–190.
114. Zhang, D.; Li, N.; Zhou, Z.H.; Chen, C.; Sun, L.; Li, S. iBAT: Detecting anomalous taxi trajectories from GPS traces. In Proceedings of the 13th International Conference on Ubiquitous Computing, Beijing, China, 17–21 September 2011; pp. 99–108.
115. Chen, C.; Zhang, D.; Castro, P.S.; Li, N.; Sun, L.; Li, S. Real-time detection of anomalous taxi trajectories from GPS traces. In *Mobile and Ubiquitous Systems: Computing, Networking, and Services*; Springer: Berlin, Germany, 2012; pp. 63–74.
116. Scott, M.S.; Dedel, K. *Assaults in and Around Bars*, 2nd ed.; Office of Community Oriented Policing Services: Washington, DC, USA, 2006.
117. Lynch, H.J.; Moorcroft, P.R. A spatiotemporal Ripley's K-function to analyze interactions between spruce budworm and fire in British Columbia, Canada. *Can. J. For. Res.* **2008**, *38*, 3112–3119.
118. Guting, R.; Schneider, M. *Moving Object Databases*; Morgan Kaufmann: Burlington, MA, USA, 2005.
119. Koubarakis, M.; Sellis, T.; Frank, A.; Grumbach, S.; Guting, R.; Jensen, C.; Lorentzos, N.; Schek, H.J.; Scholl, M. *Spatio-Temporal Databases: The Chorochronos Approach, LNCS 2520*; Springer: Berlin, Germany, 2003; Volume 9.
120. Celik, M.; Shekhar, S.; Rogers, J.P.; Shine, J.A. Mixed-drove spatiotemporal co-occurrence pattern mining. *IEEE Trans. Knowl. Data Eng.* **2008**, *20*, 1322–1335.
121. Cao, H.; Mamoulis, N.; Cheung, D.W. Mining frequent spatio-temporal sequential patterns. In Proceedings of the 5th International Conference on Data Mining (ICDM '05), New Orleans, LA, USA, 26–30 November 2005; pp. 82–89.
122. Verhein, F. *Mining Complex Spatio-Temporal Sequence Patterns*; SDM: Sparks, USA, 2009.

123. Li, Y.; Bailey, J.; Kulik, L.; Pei, J. Mining probabilistic frequent spatio-temporal sequential patterns with gap constraints from uncertain databases. In Proceedings of the 13th International Conference on Data Mining (ICDM '13), Dallas, TX, USA, 7–10 December 2013; pp. 448–457.
124. Strategic Plan for the Climate change Science Program. 2003. Available online: <http://www.climate-science.gov/Library/stratplan2003/final/ccspstratplan2003-chap9.htm> (accessed on 15 May 2015).
125. Frelich, L.E.; Reich, P.B. Will environmental changes reinforce the impact of global warming on the prairie-forest border of central north america? *Front. Ecol. Environ.* **2009**, *8*, 371–378.
126. Morens, D.M.; Folkers, G.K.; Fauci, A.S. The challenge of emerging and re-emerging infectious diseases. *Nature* **2004**, *430*, 242–249.
127. Zhang, P.; Huang, Y.; Shekhar, S.; Kumar, V. Exploiting spatial autocorrelation to efficiently process correlation-based similarity queries. In Proceedings of the 8th International Symposium on Advances in Spatial and Temporal Databases (SSTD 2003), Santorini Island, Greece, 24–27 July 2003; pp. 449–468.
128. De Almeida, C.M.; Souza, I.M.; Alves, C.D.; Pinho, C.M.D.; Pereira, M.N.; Feitosa, R.Q. Multilevel object-oriented classification of quickbird images for urban population estimates; In Proceedings of the 15th ACM International Symposium on Geographic Information Systems, Seattle, USA, 2007; p. 12.
129. Little, B.; Schucking, M.; Gartrell, B.; Chen, B.; Ross, K.; McKellip, R. High granularity remote sensing and crop production over space and time: NDVI over the growing season and prediction of cotton yields at the farm field level in Texas. In Proceedings of the IEEE International Conference on Data Mining Workshops (ICDMW '08), Pisa, Italy, 15–19 December 2008; pp. 426–435.
130. Friedl, M.A.; Brodley, C.E. Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* **1997**, *61*, 399–409.
131. Subbian, K.; Banerjee, A. *Climate Multi-Model Regression Using Spatial Smoothing*; SDM: Austin, USA, 2013.
132. Pace, R.K.; Barry, R.; Clapp, J.M.; Rodriguez, M. Spatiotemporal autoregressive models of neighborhood effects. *J. Real Estate Financ. Econ.* **1998**, *17*, 15–33.
133. Elhorst, J.P. Specification and estimation of spatial panel data models. *Int. Reg. Sci. Rev.* **2003**, *26*, 244–268.
134. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer: Dordrecht, Netherlands, 1988.
135. Levine, N. *CrimeStat 3.0: A Spatial Statistics Program for the Analysis of Crime Incident Locations*; Ned Levine & Associates: Houston, TX, USA, 2004.
136. Han, J.; Kamber, M.; Tung, A.K.H. *Spatial Clustering Methods in Data Mining: A Survey*; Taylor and Francis: London, UK, 2001.
137. Agrawal, R.; Gehrke, J.; Gunopulos, D.; Raghavan, P. *Automatic Subspace Clustering of High Dimensional Data for Data Mining Applications*; ACM: New York, NY, USA, 1998.
138. Ng, R.T.; Han, J. CLARANS: A method for clustering objects for spatial data mining. *IEEE Trans. Knowl. Data Eng.* **2002**, *14*, 1003–1016.
139. Birant, D.; Kut, A. ST-DBSCAN: An algorithm for clustering spatial-temporal data. *Data Knowl. Eng.* **2007**, *60*, 208–221.

140. Ester, M.; Kriegel, H.P.; Sander, J.; Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd* **1996**, *96*, 226–231.
141. Ankerst, M.; Breunig, M.M.; Kriegel, H.P.; Sander, J. OPTICS: Ordering points to identify the clustering structure. In *ACM Sigmod Record*; ACM: New York, NY, USA, 1999, Volume 28, pp. 49–60.
142. Wang, M.; Wang, A.; Li, A. Mining spatial-temporal clusters from geo-databases. In *Advanced Data Mining and Applications*; Springer: Berlin, Germany, 2006; pp. 263–270.
143. Zhang, T.; Ramakrishnan, R.; Livny, M. BIRCH: An Efficient Data Clustering Method for Very Large Databases. In *ACM SIGMOD Record*, Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data, Montreal, QC, Canada, 4–6 June 1996; Jagadish, H.V., Mumick, I.S., Eds.; ACM Press: New York, NY, USA, 1996; pp. 103–114.
144. Karypis, G.; Han, E.H.; Kumar, V. Chameleon: Hierarchical clustering using dynamic modeling. *IEEE Comput.* **1999**, *32*, 68–75.
145. Lee, J.G.; Han, J.; Whang, K.Y. Trajectory clustering: A partition-and-group framework. In Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data; ACM: New York, NY, USA, 2007; pp. 593–604.
146. Sander, J.; Ester, M.; Kriegel, H.P.; Xu, X. Density-based clustering in spatial databases: The algorithm gdbscan and its applications. *Data Min. Knowl. Discov.* **1998**, *2*, 169–194.
147. Lee, A.J.; Chen, Y.A.; Ip, W.C. Mining frequent trajectory patterns in spatial-temporal databases. *Inf. Sci.* **2009**, *179*, 2218–2231.
148. Chandola, V.; Kumar, V. Summarization-compressing data into an informative representation. *Knowl. Inf. Syst.* **2007**, *12*, 355–378.
149. Pan, B.; Demiryurek, U.; Banaei-Kashani, F.; Shahabi, C. Spatiotemporal summarization of traffic data streams. In Proceedings of the ACM SIGSPATIAL International Workshop on GeoStreaming, San Jose, CA, USA, 3–5 November 2010; pp. 4–10.
150. Evans, M.R.; Oliver, D.; Shekhar, S.; Harvey, F. Summarizing trajectories into k-primary corridors: A summary of results. In Proceedings of the 20th International Conference on Advances in Geographic Information Systems, Redondo Beach, CA, USA, 6–9 November 2012; pp. 454–457.
151. Evans, M.R.; Oliver, D.; Shekhar, S.; Harvey, F. Fast and Exact Network Trajectory Similarity Computation: A Case-study on Bicycle Corridor Planning. In *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing*; ACM: New York, NY, USA, 2013; pp. 9:1–9:8.
152. Kulldorff, M. SaTScan User Guide for Version 9.0; Available online: www.satscan.org (accessed on 15 May 2015)
153. Jain, A.; Murty, M.; Flynn, P. Data clustering: A review. *ACM Comput. Surv. (CSUR)* **1999**, *31*, 264–323.
154. Chang, W.; Zeng, D.; Chen, H. Prospective spatio-temporal data analysis for security informatics. In Proceedings of the 2005 IEEE Intelligent Transportation Systems, Vienna, Austria, 13–15 September 2005; pp. 1120–1124.

155. Tango, T.; Takahashi, K.; Kohriyama, K. A space-time scan statistic for detecting emerging outbreaks. *Biometrics* **2011**, *67*, 106–115.
156. Neill, D.; Moore, A.; Sabhnani, M.; Daniel, K. Detection of emerging space-time clusters. In Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery in Data Mining, Chicago, IL, USA, 21–24 August 2005; pp. 218–227.
157. Chandola, V.; Hui, D.; Gu, L.; Bhaduri, B.; Vatsavai, R. Using time series segmentation for deriving vegetation phenology indices from MODIS NDVI data. In Proceedings of the 2010 IEEE International Conference on Data Mining Workshops (ICDMW), Sydney, NSW, Australia, 13 December 2010.
158. Worboys, M.; Duckham, M. *GIS: A Computing Perspective*; CRC Press: Sacramento, CA, USA, 2004.
159. Bujor, F.; Trouvé, E.; Valet, L.; Nicolas, J.M.; Rudant, J.P. Application of log-cumulants to the detection of spatiotemporal discontinuities in multitemporal SAR images. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 2073–2084.
160. Kosugi, Y.; Sakamoto, M.; Fukunishi, M.; Lu, W.; Doihara, T.; Kakumoto, S. Urban change detection related to earthquakes using an adaptive nonlinear mapping of high-resolution images. *IEEE Geosci. Remote Sens. Lett.* **2004**, *1*, 152–156.
161. Di Martino, G.; Iodice, A.; Riccio, D.; Ruello, G. A novel approach for disaster monitoring: Fractal models and tools. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1559–1570.
162. Radke, R.; Andra, S.; Al-Kofahi, O.; Roysam, B. Image change detection algorithms: A systematic survey. *IEEE Trans. Image Process.* **2005**, *14*, 294–307.
163. Thoma, R.; Bierling, M. Motion compensating interpolation considering covered and uncovered background. *Signal Process.: Image Commun.* **1989**, *1*, 191–212.
164. Aach, T.; Kaup, A. Bayesian algorithms for adaptive change detection in image sequences using Markov random fields. *Signal Process.: Image Commun.* **1995**, *7*, 147–160.
165. Chen, G.; Hay, G.J.; Carvalho, L.M.; Wulder, M.A. Object-based change detection. *Int. J. Remote Sens.* **2012**, *33*, 4434–4457.
166. Desclee, B.; Bogaert, P.; Defourny, P. Forest change detection by statistical object-based method. *Remote Sens. Environ.* **2006**, *102*, 1–11.
167. Im, J.; Jensen, J.; Tullis, J. Object-based change detection using correlation image analysis and image segmentation. *Int. J. Remote Sens.* **2008**, *29*, 399–423.
168. Aach, T.; Kaup, A.; Mester, R. Statistical model-based change detection in moving video. *Signal Process.* **1993**, *31*, 165–180.
169. Rignot, E.J.; van Zyl, J.J. Change detection techniques for ERS-1 SAR data. *IEEE Trans. Geosci. Remote Sens.* **1993**, *31*, 896–906.
170. Im, J.; Jensen, J. A change detection model based on neighborhood correlation image analysis and decision tree classification. *Remote Sens. Environ.* **2005**, *99*, 326–340.
171. Yakimovsky, Y. Boundary and object detection in real world images. *J. ACM (JACM)* **1976**, *23*, 599–618.

172. Douglas, D.H.; Peucker, T.K. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartogr. Int. J. Geogr. Inf. Geovisualization* **1973**, *10*, 112–122.
173. Kulldorff, M.; Athas, W.; Feurer, E.; Miller, B.; Key, C. Evaluating cluster alarms: A space-time scan statistic and brain cancer in Los Alamos, New Mexico. *Am. J. Public Health* **1998**, *88*, 1377–1380.
174. Kulldorff, M. Prospective time periodic geographical disease surveillance using a scan statistic. *J. R. Stat. Soc. Ser. A (Stat. Soc.)* **2001**, *164*, 61–72.
175. Website, A. ArcGIS. Available online: <http://www.arcgis.com/>. (accessed on 15 May 2015).
176. Website, Q. QGIS: A Free and Open Source Geographic Information System. Available online: <http://www.qgis.org/>. (accessed on 15 May 2015).
177. Bivand, R. CRAN Task View: Analysis of Spatial Data. Available online: <https://cran.r-project.org/web/views/Spatial.html> (accessed 15 May 2015).
178. MathWorks. Mapping Toolbox in Matlab. Available online: <http://www.mathworks.com/products/mapping/> (accessed on 15 May 2015).
179. SAS. Spatial Analysis in SAS. Available online: <http://support.sas.com/rnd/app/stat/procedures/SpatialAnalysis.html> (accessed on 15 May 2015).
180. Beinat, E.; Godfrind, A.; Kothuri, R.V. *Pro Oracle Spatial*; Apress: New York, NY, USA, 2004.
181. Chamberlin, D. *Using the New DB2: IBM's Object Relational System*; Ap Professional: Waltham, USA, 1997.
182. PostGIS. Available online: <http://postgis.refrains.net/>. (accessed on 15 May 2015).
183. ESRI. Breathe Life into Big Data: ArcGIS Tools and Hadoop Analyze Large Data Stores. Available online: <http://www.esri.com/esriOnews/arcnews/summer13articles/breatheOlifeOintoObigOdata!> (accessed on 15 May 2015).
184. ESRI. ESRI: GIS and Mapping Software. Available online: <http://www.esri.com>. (accessed on 15 May 2015).
185. Aji, A.; Wang, F.; Vo, H.; Lee, R.; Liu, Q.; Zhang, X.; Saltz, J. Hadoop GIS: A high performance spatial data warehousing system over mapreduce. *Proc. VLDB Endow.* **2013**, *6*, 1009–1020.
186. Eldawy, A.; Mokbel, M.F. SpatialHadoop: A MapReduce framework for spatial data. In Proceedings of the IEEE International Conference on Data Engineering (ICDE'15), Seoul, Korea, 13–17 April 2015.
187. Isaak, D.J.; Peterson, E.E.; Ver Hoef, J.M.; Wenger, S.J.; Falke, J.A.; Torgersen, C.E.; Sowder, C.; Steel, E.A.; Fortin, M.J.; Jordan, C.E.; others. Applications of spatial statistical network models to stream data. *Wiley Interdiscip. Rev.: Water* **2014**, *1*, 277–294.
188. Oliver, D.; Bannur, A.; Kang, J.M.; Shekhar, S.; Boussetaire, R. A k-main routes approach to spatial network activity summarization: A summary of results. In Proceedings of the 2010 IEEE International Conference on Data Mining Workshops (ICDMW), Sydney, NSW, Australia, 13 December 2010; pp. 265–272.
189. Oliver, D.; Shekhar, S.; Kang, J.M.; Laubscher, R.; Carlan, V.; Bannur, A. A k-main routes approach to spatial network activity summarization. *IEEE Trans. Knowl. Data Eng.* **2014**, *26*, 1464–1478.

190. Gunturi, V.M.V.; Shekhar, S. Lagrangian xgraphs: A logical data-model for spatio-temporal network data: A summary. In Proceedings of the Advances in Conceptual Modeling—ER 2014 Workshops, ENMO, MoBiD, MReBA, QMMQ, SeCoGIS, WISM, and ER Demos, Atlanta, GA, USA, 27–29 October 2014; pp. 201–211.
191. Gunturi, V.M.; Nunes, E.; Yang, K.; Shekhar, S. A critical-time-point approach to all-start-time lagrangian shortest paths: A summary of results. In *Advances in Spatial and Temporal Databases*; Pfoser, D., Tao, Y., Mouratidis, K., Nascimento, M., Mokbel, M., Shekhar, S., Huang, Y., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; Volume 6849, pp. 74–91.
192. Gunturi, V.; Shekhar, S.; Yang, K. A Critical-time-point Approach to All-departure-time Lagrangian Shortest Paths. *IEEE Trans. Knowl. Data Eng.* **2015**, *27*, 2591–2603.
193. Speed, J. IoT for V2V and the Connected Car. Available online: <http://www.slideshare.net/JoeSpeed/aw-megatrends-2014-joe-speed> (accessed on May 15 2015).
194. Ali, R.Y.; Gunturi, V.M.; Kotz, A.; Shekhar, S.; Northrop, W. Discovering non-compliant window co-occurrence patterns: A summary of results. In Proceedings of the 14th International Symposium on Spatial and Temporal Databases, Seoul, South Korea, 26–28 August 2015.
195. Avery, C. Giraph: Large-scale graph processing infrastructure on hadoop. In Proceedings of the Hadoop Summit, Santa Clara, CA, USA, 29 June 2011.
196. Low, Y.; Gonzalez, J.E.; Kyrola, A.; Bickson, D.; Guestrin, C.E.; Hellerstein, J. Graphlab: A New Framework for Parallel Machine Learning. Available online: <http://arxiv.org/pdf/1408.2041>. (accessed on 15 May 2015).
197. Malewicz, G.; Austern, M.H.; Bik, A.J.; Dehnert, J.C.; Horn, I.; Leiser, N.; Czajkowski, G. Pregel: A system for large-scale graph processing. In Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data, Indianapolis, IN, USA, 6–11 June 2010; pp. 135–146.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).