International Journal of Enhanced Research in Management & Computer Applications
ISSN: 2319-7471, Vol. 6 Issue 6, June-2017, Impact Factor: 3.578

A Study of Estimating the Spatio-Temporal Patterns Prediction in Data Mining Environment

Srikanth Lakumarapu[1], Dr. Rashmi Agarwal[2]

[1]Research Scholar, Department of Computer Science & Engineering, Madhav University
[2]Associate Professor, Department of Computer Science & Engineering, Madhav University

ABSTRACT

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. 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.

Keywords: spatiotemporal data mining; survey; patio-temporal 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 database. 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.

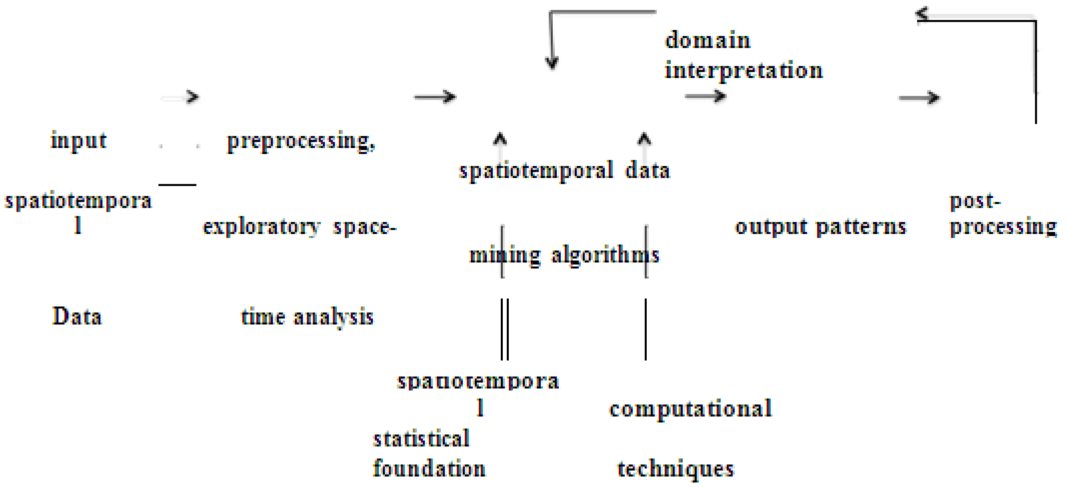

Figure 1. The process of spatiotemporal data mining.

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.

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) .

Table 1: Taxonomy of Spatial and Spatiotemporal Data Models.

Spatial Data		Temporal Snapshots (Time Series)	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 process point : 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	graph	spatiotemporal network: 1. time expanded graph,	addition or removal of nodes and edges	1. random geometric graph;

Model	time aggregated graph; 2. network flow	2. spatiotemporal event or process on spatial network;

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	nominal	Explicit
	ordinal	Arithmetic
	interval	Ordering
	Ratio	instance of
		subclass of
spatial	location	Often implicit
	area	set space: union, intersection, membership, etc.
	perimeter	topological space: meet, within, overlap, etc.
	shape	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), line(s), polygon(s)	1. spatiotemporal predicates [43] 2. trajectory distance [44,45] 3. spatial time series	1. displacement/motion speed/acceleration attraction/repulsion 2. (for line and polygon) extension/expansion, rotatio deformation	1. spatiotemporal co-variance [38] 2. spatiotemporal coupling for point events or extended spatial objects [48–

		correlation [46], tele-connection [47]	n, split/merge	,	53]
field model	Regular Irregular	1. cubic map algebra [54] 2. temporal correlation, tele-connection	local, focal, zonal change, across snapshots [29]		cellular automation [55]
spatial network	graph	1. predecessor/successor on a Lagrangian path 2. temporal centrality [56]	change in connectivity	centrality,	spatiotemporal coupling of network events

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.

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.

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.

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.

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.

4. OUTPUT PATTERN FAMILIES

Spatiotemporal Outlier
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.

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.

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].

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.

Spatiotemporal Couplings and Tele-Couplings
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.

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].

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.

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].

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 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.

Spatiotemporal Prediction
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.

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].

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-stationary, as well as the multi-scale effect (introduced in the Section 3).

Common Approaches

Spatiotemporal Autoregressive Regression (STAR): In the spatial auto regression model, the spatial dependencies of the error term, or, the dependent variable, are directly modeled in the regression equation [134]. If the dependent values yi are related to each other, then the regression equation can be modified as y = W y + X + , where W is the neighborhood relationship contiguity matrix and 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 Partitioning and Summarization
What is Spatiotemporal Partitioning and Summarization?

Spatiotemporal partitioning, or Spatio-temporal 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.

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].

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.

Table 4: Summarization Framework for Various Data Types.

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

Spatiotemporal Hotspots
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.

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].

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).

5. SPATIAL AND SPATIOTEMPORAL ANALYSIS TOOLS

This section lists currently existing spatial and spatiotemporal analysis tools, including geographic information system (GIS) software, 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.

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 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.

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.

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, RM. 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:
[30]. An inter-disciplinary survey. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 2014,.