PAPER • OPEN ACCESS

A Literature Review of Spatio-temporal Data Analysis

To cite this article: Zhenmin Dong and Chunfang Guo 2021 *J. Phys.: Conf. Ser.* **1792** 012056

View the article online for updates and enhancements.

A Literature Review of Spatio-temporal Data Analysis

Zhenmin Dong[1*], Chunfang Guo[2]

[1]School of Economics and Management, Beijing Jiaotong University, Beijing, 100044, China

[2]School of Economics and Management, Beijing Jiaotong University, Beijing, 100044, China

[*]Corresponding author's e-mail: 19120603@bjtu.edu.cn

Abstract. In the era of massive data, people are paying more and more attention to the spatio-temporal attributes of data. Many studies have begun to explore the methods of spatio-temporal analysis and apply spatio-temporal data analysis techniques to various fields. This article surveys and summarizes many previous research results. The current situation of spatio-temporal analysis is analyzed and prospected from three aspects: spatio-temporal analysis method, spatio-temporal data model and platform construction and spatio-temporal data analysis application scenarios.

1. Introduction

Spatio-temporal analysis refers to the analysis of data with time attributes and absolute and relative positions in three-dimensional space, as well as the process and methodology of analyzing these data. The ubiquitous spatial correlation in spatial data, coupled with the randomness and complexity of the time dimension, spatio-temporal data presents the characteristics of multi-dimensional, semantic, and spatio-temporal dynamic correlation. In today's society, spatio-temporal data based on spatio-temporal analysis is widely available in the fields of society, economy, population health, military, environment, ecology, geology, geography, etc. Therefore, it is necessary to study efficient mathematical analysis models based on spatio-temporal big data and massive structured, semi-structured and unstructured data storage, multi-dimensional association analysis methods and data visualization, using big data to quickly and accurately describe the representation, measurement, state and probability of the space-time pattern.

Figure 1 shows the papers published over the years under the theme of "Spatio-temporal data Analysis". This paper selects 29 representative papers, and Table 1 shows the specific composition of the papers.

Table1.Composition of papers.

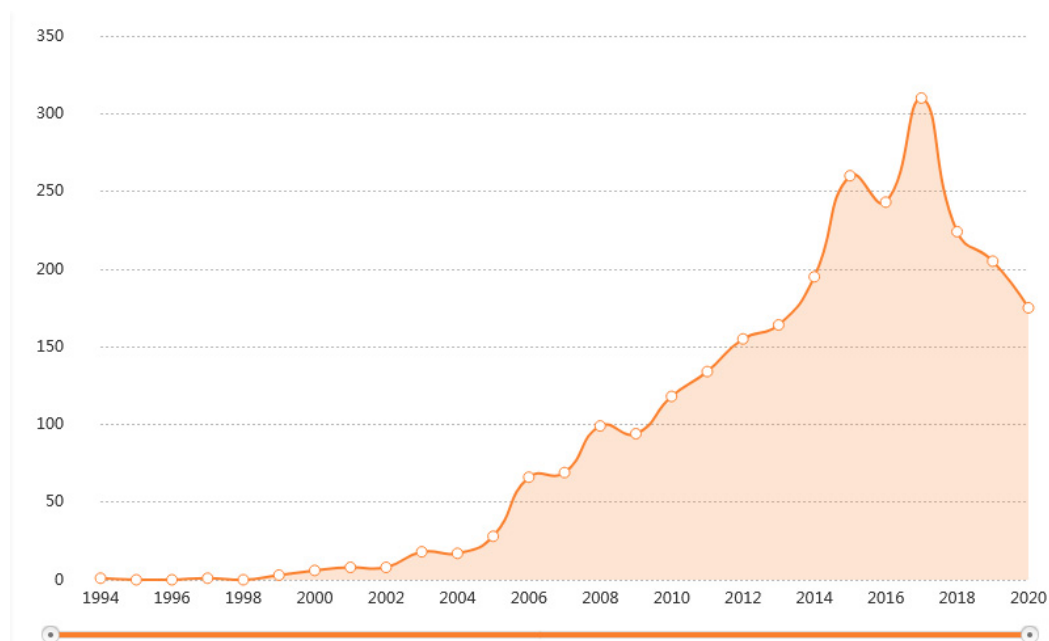| Content of papers	Number of papers	Percentage
Method research	4	13.8%
Resources construction	4	13.8%
Application research	21	72.4%

Figure 1. Number of papers published over the years.

2. Spatio-temporal analysis method

Spatio-temporal data analysis methods mainly focus on spatio-temporal data mining and clustering algorithms. Spatial clustering is an important research content of spatial data mining. [1] proposed the classification and performance requirements of spatial clustering algorithms, spatial clustering process and methods. Facing time data mining, [2] summarized research on time series clustering, including general clustering algorithms for time series clustering, standards for evaluating the performance of clustering results, and the fields of application of time series clustering. [3] proposed the process and application characteristics of clustering methods based on spatio-temporal objects, summarized the research status and main methods, characteristics and application scenarios of spatial clustering and spatiotemporal clustering, and provided ideas for spatio-temporal analysis of multi-granularity spatiotemporal objects. [4] proposed a new density-based clustering algorithm ST-DBSCAN to achieve the purpose of discovering clusters based on the non-spatial, spatial and temporal values of objects.

3. Spatio-temporal data model and platform construction

As the carrier of geospatial information for smart cities, the urban spatio-temporal information platform is an important spatial information infrastructure for smart cities. Based on this, individuals and government agencies began to devote themselves to the construction and development of spatio-temporal data models and spatio-temporal data platforms. [5] proposed a crowd-sourced spatio-temporal data model that takes into account the credibility of crowd-sourced spatio-temporal data by analyzing the interaction mechanism among target states, object versions, contributors, credibility, and events that change the state of real-space entities or information space objects in crowd-sourced spatio-temporal data. [6] proposed that different types of data content must be based on cloud platforms to realize data analysis, which analyzes the relationship between space and attributes, and the supporting relationship between status and background in the design of spatio-temporal database. It provide reference for related content of database design and research. [7] explained in detail the components of Smart Guangzhou Space-Time Big Data in terms of resource aggregation, spatial processing, data engine, management analysis, etc. With the construction of the intelligent Xuzhou spatio-temporal information cloud platform as the research background, [8] analyzed the concept, content and organizational management mode of spatio-temporal big data. In response to the fusion

problem of multi-source heterogeneous spatiotemporal information, a multi-source heterogeneous spatiotemporal information fusion service model driven by service collaboration theory is proposed.

4. spatio-temporal data analysis application scenario

There are many application scenarios for spatio-temporal data analysis. This article will explain these application scenarios from the following aspects.

4.1. Urban transportation

Taking part of the traffic flow data of Jiaonan City in Qingdao as an example, [9] used R language to visualize the traffic flow in different forms. [10] studied the time and space distribution characteristics of Beijing subway passenger flow by calculating the travel time and OD matrix of 4.31 million smart transportation card data. [11] analyzed the road section where the taxi pick-up and drop-off incidents are located, and obtained popular road sections and areas where residents travel at different times, so as to analyze the characteristics of residents' travel time and space distribution and understand the current traffic situation and residents' travel needs. Taking the taxi GPS trajectory data in the urban area of Fuzhou as the data source, [12] studied the temporal and spatial laws of residents' activities in geographic space. [13] using Xiamen's public transport IC card data and vehicle GPS data as data sources, researched the temporal and spatial distribution of public transport passenger flow. Aiming at the "on-demand" feature of shared bikes, [14] studied the sample data of shared bikes in Wuhan, Xi 'an and Hangzhou in one day by using mathematical statistics, and analyzed the distribution, spatial characteristics, cycling characteristics and turnover rate of shared bikes.

At the same time, related research has gradually begun to pay attention to traffic anomalies and accidents. [15] proposed an anomaly detection method that can be applied to intelligent transportation systems. This method is based on multi-channel singular spectrum analysis (MSSA) to characterize the temporal and spatial characteristics of transportation networks. Taking the traffic accident record data of the past ten years in Hefei City as the research object, [16] used visual analysis methods to analyze the time and location information of the accident in the traffic accident record data, explored the spatio-temporal patterns of traffic accidents, and constructed visual analysis of traffic accidents system to assist relevant departments in improving frequent traffic accidents.

4.2. Tourism

Spatio-temporal analysis is widely used in tourism and passenger flow analysis. [17] Taking the self-driving tourist hotspots in the Tibet Autonomous Region as an example, using Python to mine 6748 pieces of "footprint" data and 924 pieces of trajectory information, combined with OSM road network data and tourist attractions POI data to explore the spatio-temporal behavior characteristics of self-driving tourists. [18] used ESTDA exploratory spatio-temporal data analysis to study the tourist market and passenger flow distribution in the Beijing-Tianjin-Hebei region.

4.3. Climate environment

Spatio-temporal data can also be applied to environmental climate change analysis.[19] based on the satellite land surface temperature data from 2001 to 2012, combined with GIS technology, analyzed the day and night, seasonal and interannual changes of land surface temperature, and obtained the temporal and spatial characteristics of the heat island effect in Jinan.

In addition, more studies have begun to focus on the temporal and spatial characteristics of air pollutants.[20] used the CO data observed by MOPITT from 2000 to 2015 to show the spatial distribution and temporal changes of tropospheric CO concentration. Based on the SO_2column concentration data monitored by the Ozone Satellite Sensor (OMI) since 2008 and the SO_2 observation data (OBS) of ground monitoring stations since 2013, [21] used spatial interpolation, correlation analysis and other methods to explore the seasonal spatial distribution characteristics of atmospheric SO_2 in the Fenwei Plain, monthly changes, and average annual total changes in different periods. [22]

predicted the concentration of six standard air pollutants through time and space modeling of particulate matter and gaseous pollutants in Beijing.

At the same time, the smog problem has also been a research hotspot in recent years. Except for Beijing, Tianjin, Hebei and other heavily polluted areas, due to limited research on the temporal and spatial accumulation of PM2.5 concentration, there is little discussion about regional integration of air quality management in other parts of China. In order to fill this gap, [23] conducted temporal and spatial clustering of PM2.5 concentration at the level of Chinese cities, and studied the spatial self-aggregation effect of PM2.5 concentration in Chinese cities.

[24] based on the temporal and spatial changes of man-made marine debris on Chilean beaches, tested the hypothesis that the amount of man-made marine debris (AMD) on the beaches increased over time in emerging economies such as Chile. [25] studied the temporal and spatial changes of chemical substances in the coastal marine environment of Antarctica, the proposed modeling method can provide insight into this behavior and supplement physical sampling that is difficult to perform in this environment.

4.4. Population distribution and user behavior
Spatio-temporal data can be used to measure population density. [26] chose kindergarten, a pre-school education facility that can represent population agglomeration, and analyzed the spatial distribution of kindergartens in Urumqi through the standard deviation ellipse, average nearest neighbor, and kernel density. And it used this as a basis to explore the population agglomeration status in various regions of Urumqi;

User behavior is also the focus of many studies. [27] proposed that indoor positioning data records the spatio-temporal trajectory of users in indoor space activities, and it is an important information source for studying indoor behavior of people. It faces the characteristics of time and space distribution of indoor people, builds a visual analysis model of time and space behavior and realizes a shopping mall customer flow analysis system oriented to shopping malls;

In addition, there are studies starting from mobile phone data, relying on mobile phone call location data in a megacity in central my country, and statistically analyzing and spatially visualizing thespatio-temporal behavior characteristics of urban residents.

At the same time, spatio-temporal data mining on social media location service big data can provide decision-making basis for applications such as urban planning, business decision-making, and user behavior analysis. [28] based on the Sina Weibo sign-in point data, conducted a spatio-temporal interaction test, determined an appropriate spatio-temporal analysis scale, and used spatio-temporal rearrangement scanning statistical methods to mine spatio-temporal hot spots on a short time scale and a long time scale respectively. [29] used emerging crowdsourcing data to reflect people's mobility and aggregation with social media check-in within 24 hours, and proposed a combination of kernel density estimation (KDE), geographic and time weighted regression (GTWR) and the framework of the Herfindahl-Hirschman Index (HHI) to explore the temporal and spatial distribution characteristics of vibration and the temporal and spatial relationship with influencing factors.

5. Conclusion
This article summarizes the analysis methods, data models and typical application scenarios of spatiotemporal data, and provides references for the application of spatiotemporal data analysis in other fields.

References
[1]   Liu S,Ji GL. (2010) A Review of Researches on Spatial Clustering.JOURNAL OF NANJING NORMAL UNIVERSITY(ENGINEERING AND TECHNOLOGY EDITION), 10(2):57-62.
[2]   Liao T W. (2005) Clustering of time series data—a survey. Pattern Recognition, 38(11):1857-1874.

[3]   Yang Z K, Li X, Chen D. (2018) Study on Clustering Methods and Applications for Spatio-temporal Objects.GEOMATICS WORLD,25(2):40-44.

[4]   Birant D, Kut A. (2007) ST-DBSCAN: An algorithm for clustering spatial–temporal data. Data & Knowledge Engineering, 60(1):208-221.

[5]   Zhou X G, Zhao Y J, Li G Q, et al. (2018) A multi-source spatio-temporal data model that takes account of credibility.Journal of Wuhan University (Information Science edition),43(1):10-16.

[6]   Wang T M, Li Y,Liang J P.(2017) Design and research of spatio-temporal database of smart city.Beijing surveying and mapping,(6):72-76.

[7]   Zhang P C,He H G,Yang W J, et al. (2018) Design and Construction of Smart Guangzhou Spatio－temporal Big Data Framework. GEOMATICS & SPATIAL INFORMATION TECHNOLOGY,41(5):11-13,17.

[8]   Huang Q, Wang Y J, Chu L, et al. (2018) Intelligent Xuzhou Spatio-temporal information Service Model Research. GEOSPATIAL INFORMATION,16(5):31-34.

[9]   Li X W, Wang H Q. (2019) Visualization of traffic flow data based on R language. GEOSPATIAL INFORMATION, 017(004):95-102.

[10] Huang J, Wang J E, Jin H T, et al. (2018) Spatial and temporal distribution pattern and Characteristics of Subway passenger flow in Beijing -- Based on intelligent traffic card data.Progress in Geography,37(3):397-406.

[11] Han Y T, Wang W A. (2018) On the Spatial and Temporal Distribution of Resident Trip Based on Taxi GPS Data.GEOMATICS & SPATIAL INFORMATION TECHNOLOGY,41(2):87-89.

[12] Liu M,Wu Q Y.(2018) Spatial and temporal patterns of residents activity intensity based on taxi OD data.Journal of Fuzhou University(Natural Science Edition),46(2):204-209.

[13] Wu Q Y,Su K Y,Zou Z J. (2018) Spatio-temporal analysis of bus passenger flow based on massive IC card data.Journal of Guizhou University(Natural Sciences),35(6):93-98,105.

[14] Zhou R,Wang Y Q,Zhu L, et al. (2019) Research on the characteristics of bike-sharing based on temporal and spatial data.Journal of Wuhan University of Technology Transportation Science & Engineering),43(1):159-163.

[15] Mai H. Hassan,Ali Tizghadam,Alberto Leon-Garcia. (2019) Spatio-temporal Anomaly Detection in Intelligent Transportation Systems. Procedia Computer Science,151.

[16] Rao Y M,Zhang Y K,Xie W J, et al.(2019) Visual analysis Method of traffic accident spatial-temporal patterns.COMPUTER SCIENCE,46(4):14-21.

[17] Liu Y P, Bao J G, Huang Y H, et al.(2019) Research on Spatio-temporal Behavior of Self-driving Tourists based on GPS Data -- A case study of Xizang Province. WORLD REGIONAL STUDIES, 2019, 28(01):149-160.

[18] Xu Y, Tao Z M, Ji X M, et al.(2018) Spatial- Temporal Evolution Characteristics of Tourism Flow in Beijing－Tianjin－Hebei Area. JOURNAL OF NANJING NORMAL UNIVERSITY(Natural Science Edition).

[19] Sun Z D,Lian L S,Zhang K X, et al. (2019) Spatial and temporal characteristics of urban heat island effect in Jinan city based on MODIS data.Journal of Qufu Normal University,45(1):86-94.

[20] Zhao X L,Chi T H. (2018) Spatial distribution and temporal variation of tropospheric CO concentration over China based on MOPITT measurements [J].Environmental Science & Technolog,41(4):71-76.

[21] Wei W, Wang L J, Jin Z H, et al.(2018) The spatio-temporal distribution characteristics of atmospheric SO_2 in Fenwei Plain based on OMI data. Ecology and Environmental Sciences, 027(012):2276-2283.

[22] Xu J, Yang W, Han B, et al. (2019) An advanced spatio-temporal model for particulate matter and gaseous pollutants in Beijing, China. Atmospheric environment, 211(AUG.):120-127.

[23] Ziyue, Chen, et al. (2019) Spatial self-aggregation effects and national division of city-level PM2.5 concentrations in China based on spatio-temporal clustering. Journal of Cleaner Production.

[24] Hidalgo-Ruz V, Honorato-Zimmer D, Gatta-Rosemary M, et al. (2018) Spatio-temporal variation of anthropogenic marine debris on Chilean beaches. Marine Pollution Bulletin, 126(JAN.):516-524.

[25] Michael L Bates, Darryl W Hawker, Roger Cropp. (2019) Spatio-temporal variations in hexachlorobenzene partitioning in a near shore Antarctic marine environment from a one-dimensional coupled ecosystem-chemical distribution model. 196:65-76.

[26] Jiang Y, Dou Y, Wang Z H. (2018) Spatial-temporal evolution analysis of population distribution in Urumqi based on POI data.Statistics and Management,(11):78-82.

[27] Cheng D Y, Qin K, Pei T, et al.Visual analysis design and implementation for group spatiotemporal behavior based on indoor position data.Journal of Geo-information Science,2019,21(1):36-45.

[28] Xu Z Y, Xiong Y, Gao R G. (2018) Mining spatial-temporal hot spots with microblog check-in data-A case study of Beijing.Engineering of Surveying and Mapping,27(5):10-16.

[29] Wu C, Ye X, Ren F, et al. (2018) Check-in behaviour and spatio-temporal vibrancy: An exploratory analysis in Shenzhen, China[J]. Cities, 77(jul.):104-116.