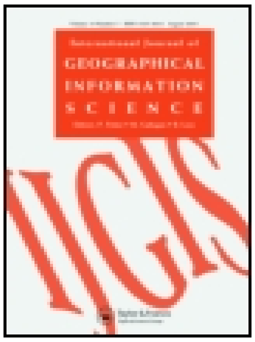

International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: https://www.tandfonline.com/loi/tgis20

Big Spatiotemporal Data Analytics: a research and innovation frontier

Chaowei Yang, Keith Clarke, Shashi Shekhar & C. Vincent Tao

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
https://doi.org/10.1080/13658816.2019.1698743

Taylor & Francis
Taylor & Francis Group

EDITORIAL

Check for updates

Big Spatiotemporal Data Analytics: a research and innovation frontier

1. Introduction

Big Data have emerged and become the norm in the past decade with its well-known 4V challenges and now adds value (a 5th V) for scientific research and the development of new applications. Such Big Data are normally used to record phenomena of interest from social, physical and economical perspectives. Our 4-dimensional (4D) world with 3D in space and 1D in time puts a stamp (or a series of stamps) on the data collected, and a spatiotemporal signature is recorded within most Big Data. We refer to such Big Data with spatiotemporal stamps as Big Spatiotemporal Data. Big Spatiotemporal Data Analytics is the study and application of thinking, algorithms, frameworks, tools, and solutions for the processing of Big Spatiotemporal Data. Most Big Data are produced with space and time stamps and are samples in sequential observations from various remote, in-situ, mobile and human sensing systems or simulations. The data can be used to study a variety of phenomena from global to microscale, for example, how global climate change in the past centuries has caused the rise of sea level and temperature, and in turn, brought sometimes catastrophic hazards. At the micro-scale, the study of DNA and cell evolution has unveil the patterns, vulnerabilities and potential cures for cancer and other diseases.

In the past decade, Big Spatiotemporal Data have driven and enabled innovations in all aspects of information systems from hardware, algorithms, software/tools, to applications and fostered the integration of different traditional disciplines to enable new research directions. Distinct from traditional data analytics, Big Spatiotemporal Data Analytics demands new frameworks and information attributes (Karim *et al*. 2018) to obtain results more efficiently when discovering trends and patterns in various domains from human dynamics (Fang *et al*. 2017), traffic congestion (He *et al*. 2017), smart cities (Machado *et al*. 2019), industry evolution (Li *et al*. 2018), medical and health issues (Kraemer *et al*. 2018), to brain science (Bassett and Sporns 2017).

To capture the latest advancements in this important research theme, we organized this special section and conducted a brief review. We searched for and analyzed publications using the keywords 'spatiotemporal' and 'Big Data' in the Web of Science with the intention of identifying only space-time integrated Big Data research. The search found approximately 300 papers, and we further analyzed the temporality of the publications and the collaborations among different countries. Section 2 presents a statistical analysis of the literature identified. Section 3 introduces the papers selected in the context of scholar contributions; Section 4 analyzes the broad application domains, and Section 5 discusses the future research and innovation directions. We conclude with a vision for future research and broad impact in section 6.

2. Literature analyses

Scholars in most academically active countries have published on Big Spatiotemporal Data in the past decade (Figure 1). The first paper on Big Spatiotemporal Data appeared in 2009, which implies a 10-year history of scholarship in this field with a rapidly increasing upward trajectory (Figure 2). Most publications were from the USA and China, but other

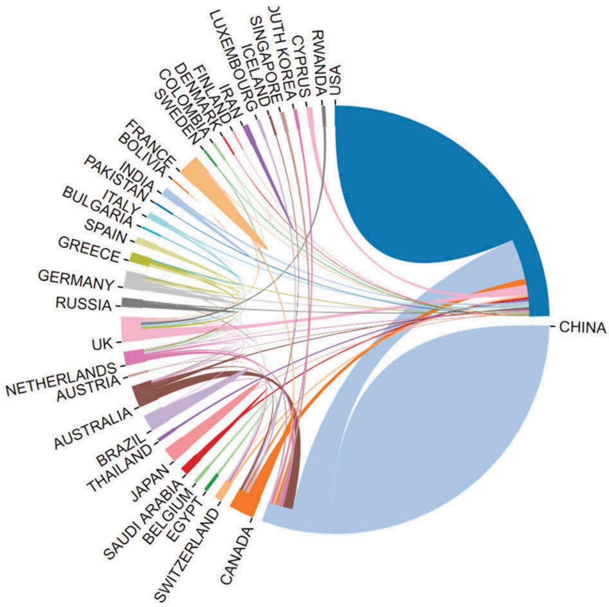

Figure 1. The Big Spatiotemporal Data Analytics publication distribution and collaborations among different countries (as of 25 Nov 2019).

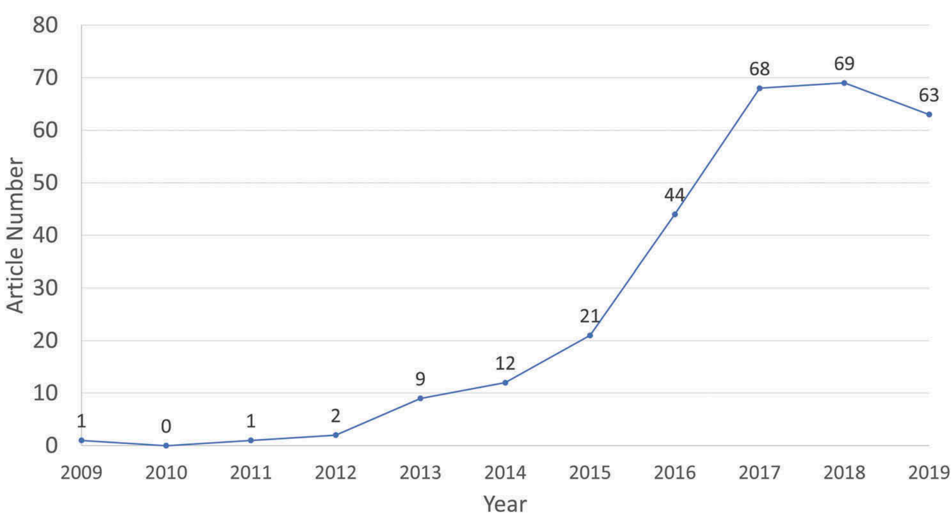

Figure 2. The past decade scholarship landscape and trajectory in the Big Spatiotemporal Data area (as of 25 Nov 2019).

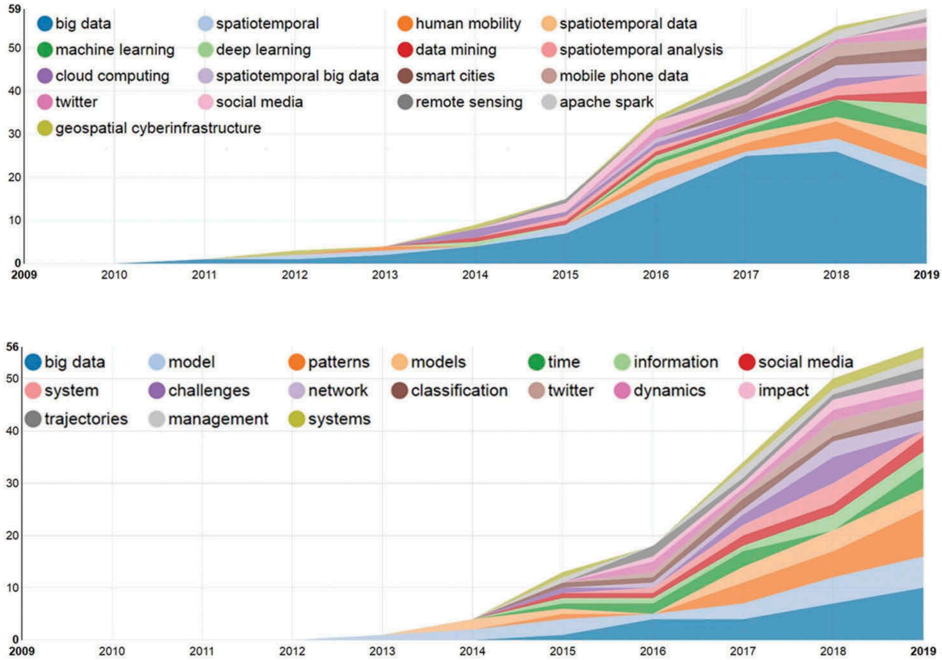

Figure 3. The emerging topics taken from keywords (upper) and extended keywords (lower) for Big Spatiotemporal Data research (as of 25 Nov 2019).

developed countries such as Canada, Germany, Japan, Australia and France also contributed significantly. Collaborations among these countries remain very active (Figure 1).

We further analyzed the emerging concepts of special interest and identified the following as top keywords in the publications (Figure 3), including those related to human dynamics (e.g. human mobility, social media, mobile phone data, twitter), technology and methods (e.g. spatiotemporal analyses, data mining, machine learning, deep learning, cloud computing, Hadoop/Spark, network), and data and information (e.g. data, metadata, models, patterns, trajectories, systems). The literature review affirms that spatiotemporal Big Data analytics is a flourish new research direction with a broad impact across different disciplines.

3. The special section

Through a rigorous review process, seven papers were selected that represent a variety of disciplines in landuse, climate change, transportation, human mobility, logistics flow, air quality and nighttime studies. These studies applied diverse datasets from both physical and social domains. Together, these papers contributed new algorithms, frameworks, approaches, and solutions to address specific domain challenges and pushed the advancement of Big Spatiotemporal Data research. The special section includes the following papers:

- To address the labor-intensive and cost challenges of landuse characterization at the urban-object level, Srivastavaa et al. (2018) proposed a deep learning solution to automatically mine and predict the fine-grained landuse classes from globally available open data, such as Google Street View. The solution includes a convolutional neural network with shared layers among different branches of the network and a variable number of pictures as input. They validated the results using OpenStreetMap open data and showed better results than those using other deep-learning methods.
- Zhang et al. (2019) proposed a two-stage transportation analysis approach that reconstructed individual passenger trips and then clustered the trips using shared-flow clustering to identify the transit corridors from smart card and bus trajectory data. The research also analyzed the spatiotemporal distribution of such corridors and the results can be utilized for transit planning and management.
- Using approximately 10 million taxi trip records in New York and Shenzhen, Gong et al. (2019) proposed a two-layer framework to extract activity patterns for travel-purpose prediction and to identify return trips and subsequent activities. They designated the first layer as an activity inference model (AIM) and the second as a pairing journey model (PJM) for the two purposes in series. The framework integrate spatiotemporal clustering, Bayesian probability and Monte Carlo simulations.
- He et al. (2019b) proposed a complex pattern mining algorithm to discover event-based spatiotemporal association patterns, representative of geographic dynamics. The algorithm adopts a hierarchical framework to support mining beyond point data representation, reveal dynamic characteristics of complex geographic phenomena and discover their associated factors. Using air quality data in the Beijing–Tianjin–Hebei region, the authors extracted the air-pollution distribution patterns and contributing factors, such as wind, temperature and humidity.
- To support hundreds of terabytes of big climate data, Li et al. (2019) proposed a scalable online visual analytics system for knowledge discovery in a scalable and shareable environment. The authors showed three advantages of the proposed framework: (a) balanced usability and flexibility for climate scientists and other stakeholders; (b) scalablility and high performance for parallel and swamping in-memory processing of hundreds of terabytes of climate data; (c) the system is sharable and reusable for different scientists to easily duplicate the analytical process and results.
- Based on 10 million orders placed in a 3-year time frame in Hong Kong, Zhao et al. (2018) conducted three analyses within a logistics-flow framework: (a) spatial analyses identified the displacement and direction of goods flows and found the pattern distinctive from human mobility in an urban setting; (b) network analysis investigated the destination and origin and found that the urban area was increasingly connected with goods shipments and (c) spatial interaction analysis revealed that distance was not influential to the spatial distribution of goods movement.
- Using approximately 80 million telecommunication records from the city of Seoul, South Korea, Kim (2019) analyzed the spatiotemporal patterns of nighttime vitality of three different city centers within the city. Functional Principal Component Analysis

(FPCA) was used to analyze the data and revealed the dynamics of time span of nighttime, the seasonal variations and regional variations in nighttime.

It is notable that although many datasets used are in the open domain, some datasets (such as data used in Zhao et al. and Kim) are sensitive and demand preprocessing to assure privacy and mitigate other sensitivity issues.

4. Big Spatiotemporal Data Analytics in action

In addition to the seven papers selected for this special section, the examination of the past-decade of publications showed extensive research on Big Spatiotemporal Data Analytics by developing new algorithms, frameworks and approaches.

Spatiotemporal data streams from mobile devices, in-situ sensing, and the Internet of Things (IoT) have been used to analyze communication traffic flow patterns for *engineering design and planning communication networks*. The spatiotemporal patterns identified have been used to, for example, optimize the IoT and cellular network deployment using limited electronic bandwidth effectively (Bothe *et al*. 2019), detect violations of electromagnetic usage (Liu *et al*. 2019), optimize and improve error detection (Gao *et al*. 2018), compress IoT data for fast transmission through limited bandwidth to enable real-time video and audio applications (Moon *et al*. 2017), and to increase the reliability of a 5G wireless sensor network. The context of these communication networks includes highly spatiotemporal dynamical systems, such as urban areas, natural environments, the battlefield and human health (Chen and Yang 2016, Machado *et al*. 2019). These and future spatiotemporal optimizations improve the infrastructure for a variety of applications that rely increasingly on mobile, IoT and wireless connections.

Urbanization is one of the most significant mass human migrations in the past centuries and has brought many challenges, for example, for crime reduction, public health, public security and planning (Gong *et al*. 2012). Spatiotemporal patterns can be mined from Big Data (social media, phone locations, bus lines and traffic zones) collected within cities and used to better understand our urban settings and broader patterns such as human mobility and accessibility (Pappalardo and Simini 2018). This is true also for population dynamics (Liu *et al*. 2018a), crash indices (Bao *et al*. 2019), poverty distribution (Njuguna and McSharry 2017), segregation based on environment, gender, racial/ethnic and socioeconomic aspects (Park and Kwan 2017), patterns of social activities for policy development (Fu *et al*. 2018), and the stability of urban human convergence and divergence patterns (Fang *et al*. 2017). For crime reduction and public safety, Big Data (such as surveillance videos) have been used to improve data service sustainability and to analyze crime patterns based on city to city distance and connectivity (Kotevska *et al*. 2017). Big spatiotemporal data have been used to detect patterns of surveillance in urban settings (Xu *et al*. 2016), to identify objects or people in real time for video surveillance (Zhang *et al*. 2016), and to provide robust and accurate identification and extraction of objects and events of interest (Xu *et al*. 2015b).

Another area of research in Big Spatiotemporal Data Analytics relates to *transportation* studies. Big spatiotemporal data, such as probe and social media data, have been explored to: provide real-time solutions for strategic air traffic management (Xie *et al*. 2019a); analyze individual trip patterns, reveal fuel-consumption and driving habits (Kan *et al*. 2018); predict road congestion patterns (He *et al*. 2017); model residents' travel patterns (Gong *et al*. 2016); and predict future traffic by matching a current situation to the most effective prediction model trained using historical data (Xu *et al*. 2015a). Open or derived data are especially useful, e.g. bullet train data in China were used to analyze travel behavior and station capability (Wei *et al*. 2017).

The worsening *air quality* has become one of the most critical urban concerns in the past decades, especially in Asia (Fuzzi *et al*. 2015). Chen *et al*. (2019) utilized bike sensor data, satellite observations and other data to analyze the hourly, daily and seasonal patterns of PM2.5/10 and their relationship to different atmospheric conditions. Antonić *et al*. (2016) utilized cloud cover data, and published and subscribed to middleware services to process mobile sensor data for air quality monitoring. The Radial Basis Function, a real value approximation assignment function, was used to conduct spatio-temporal interpolation for assessing the trends of PM2.5 air pollution (Losser *et al*. 2014). The air quality study effort was also coupled with atmospheric phenomena, such as clouds and the movement of particles on regional and global scales (Fuzzi *et al*., 2015).

The rapid movement of populations has also brought many challenges to *infectious disease* control and other *public health* problems (Frumkin 2016). Mobile, health sensor, in-situ and IoT data, in combination with satellite observations, have provided data to analyze disease (e.g. Ebola and Influenza), to model disease spread and evolution (Chen *et al*. 2017) and to model the different forms of infectious disease transmission (Kraemer *et al*. 2018). Research has also been conducted to understand the behaviors of vectors that cause disease spread, such as mosquitos. For example, in-situ sensor data were integrated with historical data in a national and international open access repository of mosquito surveillance data to understand Malaria vector population dynamics (Rund *et al*. 2019).

Big Spatiotemporal Data have also been used in oceanographic, ecosystem, climate change, and other *Earth system sciences*. Big climatic, socioeconomic and other spatio-temporal data have been analyzed to identify the correlation among grassland productivity, climate and socioeconomic variables (Xie *et al*. 2019b), to identify and predict forest loss (Harris *et al*. 2017), and to understand, predict and manage complex dynamics across different geoscience domains in a human-centered machine learning strategy, which ensures machines to continue learning from human and be designed to interact with human in an effective fashion (Peters *et al*. 2018, Crandall *et al*. 2018). Big Data demands have led to building global archives of essential biodiversity variables (Kissling *et al*. 2018). Historical Big Data collected from various sources including citizen sensor networks, such as Inaturalist, are used in the studies of zoology, botany, biodiversity, tsunamis, earthquakes, climate change and other geosciences (McKinley *et al*. 2017).

In the same context, Big Spatiotemporal Data have been used to improve the understanding of *natural hazards*, including floods, earthquakes, landslides, volcanic eruptions and tsunamis (Martinez-Alvarez and Morales-Esteban 2019). The Big Data have been used to analyze the recovery zones (such as hurricanes Sandy and Irene in New York City, Zhu *et al*. 2016), to investigate the pre-, peri- and post-stages of natural hazards for compliance measurement of responses/evacuation (Martin, Li, and Cutter 2017), and to optimize the resource deployment for hazard mitigation based on victim activity patterns mined from social media data, images and in-situ sensor observations.

5. Future directions

Big Spatiotemporal Data present a challenge to automate the extraction of actionable information and knowledge. The need for more accurate and reliable understanding and predictions requires improvements to sensor and system architecture, analytical methods, and data processing tools. The challenge still prevents Big Spatiotemporal Data Analytics from reaching its full potential (Yang *et al*. 2019).

Improvements to Big Spatiotemporal Data science requires improvements in *rules and thinking*. Spatiotemporal thinking (Yang *et al*. 2011) integrating interdisciplinary knowledge through a spatiotemporal framework is important to harness the wealth of cross-domain knowledge and to fully understand Big Spatiotemporal Data (Yang *et al*. 2017). New thinking uncovering hidden rules and new frameworks are needed to devise new methods so that patterns embedded can be revealed within improved conceptual frameworks (Karim *et al*. 2018). Progress are being made, for example developing new methods, to analyze Big Spatiotemporal Data using machine learning for natural hazards (Martinez-Alvarez and Morales-Esteban 2019). Efficiently integrating Big Data from different spatiotemporal scales is critical to Big Earth observation data for earth system sciences (Hu *et al*. 2018). Methods with spatiotemporal segmentation, feature selection, texture and fractal spectrum analysis, object recognition and semantic understanding are being developed to find patterns from image, audio, video and other multimedia data for surveillance (Xu *et al*. 2016). In addition, research are also conducted on accurate identification of objects and events of interest (Xu *et al*. 2015b) for real-time crime fighting (Zhang *et al*. 2016). Spatiotemporal proximity analyses are also being devised for many physical and social studies (Fang *et al*. 2017).

Big Spatiotemporal Data, especially open and social media data, are *heterogeneous* in their spatial, temporal, and semantic formats across resolutions, subjects and sources. There is an increased need for cross-scale data fusion, including integration across various sources and interpolation across spatiotemporal domains. New models and frameworks are required for integrating data to study air pollution, vector-borne diseases, ecological and climate relationships, and hydrological systems. For example, weighted- and unmixing-based as well as sparse representation can be used to fuse data with different spatiotemporal resolutions from the Landsat-7 and MODIS satellite observations (Wei *et al*. 2018) for ecological and land use studies.

Big Spatiotemporal Data are obtained from various sources and archived, processed at different computing modes, such as cloud computing, mobile computing, edging computing and wearable computing. Optimizing and intgratively utilizing such *Hybrid computing* modes are critical to release the full power of Big Spatiotemporal Data Analytics. For example, how to best leverage cloud computing based on the collocation of the data and computing power is critical to support large scale scientific research and engineering development, such as global land cover study (Liu *et al*. 2018b) and network simulations (Iqbal *et al*. 2019). Continuously equipping IoT devices with ubiquitous and intelligent computing capability would help improve the timeliness in detection and rapid response (Al-Turjman 2019). Increasingly sharing and empowering mobile computing with smart (health) devices would help both increasing monitoring capability and customizing services to individuals as needed (Williams *et al*. 2018).

New and interdisciplinary algorithms for the analysis of moving objects are needed for various domains such as disease spread, forecasting social events and human dynamics (Fang *et al*. 2017) because the increasing complexity of available data types and interdisciplinary engagements. Traditional Big Data tools are often not capable of supporting Big Spatiotemporal Data. For example, research has expanded Hadoop with capabilities of spatial indexing, spatiotemporal indexing (Li *et al*. 2017b), and trajectory analytics (Bakli *et al*. 2019). New machining-learning algorithms such as clustering, feature selection and extraction methods that can efficiently handle spatiotemporal and factor connections for Big Spatiotemporal Data are needed to analyze the data generated from social media and citizen sensing data, such as taxi and public transit. Further needs include methods that use these data to identify public threat events and patterns (Tang *et al*. 2019) and to forecast urban traffic patterns (Pavlyuk 2019) in a timely and accurate fashion. Clustering trajectories using GPUs has been shown to be critical to speed up computation and improve data mining capabilities in Big Spatiotemporal Data, such as for video and audio sensors (Deng *et al*. 2015). Mining patterns, information and knowledge automatically from Big Spatiotemporal Data may require us to 'reinvent the wheel', that is to adapt known algorithms and redesign or invent new ones that can utilize patterns and principles of different spatiotemporal contexts (Shekhar *et al*. 2015).

In terms of *models*, Bayes' rules and Monte Carlo simulations have been revised with spatiotemporal focus to develop models of residents' travel patterns (Gong *et al*. 2016). A prediction model was proposed to learn from the current traffic situation (or context) in real-time and predict future traffic by matching the current situation to the most similar historical data (Xu *et al*. 2015a). A multilevel model of meme diffusion was developed to illustrate the applications of Influenza outbreaks, the Ebola epidemic and marijuana legalization (Yang *et al*. 2016). As the *nature* collection of *multidisciplinary nature of machine learning* indicates that the integration of domain knowledge with models needs active research for processing Big Spatiotemporal Data for new domain knowledge and methodology (https://www.nature.com/collections/csgqqsrfxh).

Big spatiotemporal data analysis now draws on new analytical *frameworks*. For example, deep learning can be expanded to a spatiotemporal framework for different scales of problems in space and time in the recognition of citywide crash patterns (Bao, Liu, and Ukkusuri 2019). The integration of cyber-physical systems requires new frameworks and models to provide integrated management, collaborative observation, scalable processing/fusion and intelligent services (Zhang et al., 2018). For example, a systematic workflow management framework was developed for complex and real time event-driven applications in transportation and air pollution to integrate in-situ sensors and virtual simulations automatically (Alvarez, Morales and Kraak 2019).

However, to reach a broader impact and improve ease of use, the algorithms and methods for big spatiotemporal data should be implemented as new or extensions of *commercial and open-source software* (Wang *et al*. 2019). Visualization tools developed in the past are mostly for a single purpose, but a more flexible visualization approach is needed for Big Spatiotemporal Data for event visualization and the exploration of complex phenomena (He *et al*. 2019a). For example, spatiotemporal algorithms should be further developed into versatile software to handle various analytics of climate data (Li *et al*. 2017a), Earth observations (Xia *et al*. 2018), social media, real-time IoT data and others. Such software should be able to reduce storage requirements, improve query/access performance, and facilitate more convenient analytics for the broad spatiotemporal domains.

An integrative understanding of space and time is essential for *policies* in smart cities, in communities and our daily lives (Kitchin 2019), especially related to data sharing, geospatial privacy and sensitive data protection. Integrating cutting-edge technologies such as geospatial artificial intelligence, high performance computing and Graphics Processing Unit processing is critical to handle Big Spatiotemporal Data to support new software tools for revealing sensitive patterns of value to policy formulation, such as transit usage (Aqib *et al.* 2019). Utilizing crowd-sourced data involves issues of data quality, noise removal and privacy, but are valuable in many different domains of human and physical sciences, such as in the study of urbanization and climate change. Early spatiotemporal context aware and privacy aware frameworks were investigated to provide secure and private access to multimedia Big Data (Samuel *et al.* 2015), but this remains a topic of research interest until mature solutions were developed.

6. Towards a new generation geographical information system and sciences

The continuous innovation in response to the demands of Big Spatiotemporal Data Analytics not only leverages the advancements of many relevant innovations, such as cloud computing and AI (Yang *et al.* 2017), but also serves as an innovation collision point by providing abundant data to enable new scientific discoveries and next-generation engineering applications, therefore, it cultivates a new generation of GIS. In addition to driving new thinking, algorithms, models, frameworks, and software, the research on Big Spatiotemporal Data Analytics are also enabling and advancing progress on the many grand engineering and application challenges of the twenty-first century that will change our lives fundamentally. For example,

- Completely safe *driverless cars* will only be technically possible if we integrate and mine a variety of static, dynamic, real-time data by considering all environment driving conditions (such as terrain, roads, rainy, pedestrians, traffic hours) through deep and continuous data analysis (Schimbinschi *et al.* 2017).
- Similarly, the integration and analysis of industrial Big Spatiotemporal Data could help connect to logistics flow and implement predictive maintenance and reduce energy use in industrial and manufacturing facilities in the industry 4.0 era for *smart manufacturing* (Yan *et al.* 2017).
- Spatiotemporal contexts, patterns, or principles extracted from data, usage, sensors and computing are critical to ensure sustainable geographical information services for massive accesses to web mapping-based routing (Li *et al.* 2018), for VR/AR enabled situation awareness (Chen *et al.* 2018), and for decision-making support tools for *smart cities, transparent governance and critical operations*.
- The evolution of the IoT has brought machine intelligence to devices in our daily-life from cars, beds, refrigerators, TVs, to office equipments. A *smart living environment* will be based on the sustainable, reliable and trustworthy operations of the IoT devices relying on real-time and accurate analyses of the data generated by them using Geospatial AI (GeoAI) or AI for IoT (Al-Turjman 2019). Lastly,

- *Open systems and solutions* can help us integrate data that were previously not possible to understand, manage, and discover new patterns of urbanization (Liu *et al*. 2015), human dynamics (Cao *et al*. 2015), sustainable development (Xu *et al*. 2015), and of public and individual health (Williams *et al*. 2018) to support decisions for better, peaceful and more prosperous lives.

Acknowledgments

Professor May Yuan, the editor in chief of IJGIS provided continuous guidance and support to the maturity of this special section. Editing activities of Chaowei Yang and Keith Clarke were partially supported by the NSF Spatiotemporal Innovation Center. We thank all authors who submitted their manuscripts to the special section.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Science Foundation [1841520 and 1835507];

References

Al-Turjman, F., ed.. 2019. *Artificial intelligence in IoT*. New York, NY: Springer.

Antonić, A., *et al*., 2016. A mobile crowd sensing ecosystem enabled by CUPUS: cloud-based publish/ subscribe middleware for the Internet of Things. *Future Generation Computer Systems*, 56, 607–622. doi:10.1016/j.future.2015.08.005

Aqib, M., *et al*., 2019. Rapid transit systems: smarter urban planning using big data, in-memory computing, deep learning, and GPUs. *Sustainability*, 11 (10), 2736. doi:10.3390/su11102736

Bakli, M., Sakr, M., and Soliman, T.H.A., 2019. HadoopTrajectory: a Hadoop spatiotemporal data processing extension. *Journal of Geographical Systems*, 21 (2), 211–235. doi:10.1007/s10109-019-00292-4

Bao, J., Liu, P., and Ukkusuri, S.V., 2019. A spatiotemporal deep learning approach for citywide short-term crash risk prediction with multi-source data. *Accident Analysis & Prevention*, 122, 239–254. doi:10.1016/j.aap.2018.10.015

Bassett, D.S., and Sporns, O., 2017. Network Neuroscience. *Nature Neuroscience*, 20(3), 353.

Bothe, S., Qureshi, H.N., and Imran, A., 2019. Which statistical distribution best characterizes modern cellular traffic and what factors could predict its spatiotemporal variability? *IEEE Communications Letters*, 23 (5), 810–813. doi:10.1109/COML.4234

Cao, G., *et al*., 2015. A scalable framework for spatiotemporal analysis of location-based social media data. *Computers, Environment and Urban Systems*, 51, 70–82. doi:10.1016/j.compenvurbsys.2015.01.002

Chen, J.Y., *et al*., 2018. Situation awareness-based agent transparency and human-autonomy teaming effectiveness. *Theoretical Issues in Ergonomics Science*, 19 (3), 259–282. doi:10.1080/1463922X.2017.1315750

Chen, S., *et al*., 2019. A big data analysis of PM2. 5 and PM10 from low cost air quality sensors near traffic areas. *Aerosol and Air Quality Research*, 19 (8), 1721–1733. doi:10.4209/aaqr.2019.06.0328

Chen, Y., *et al*., 2017. Reality mining: a prediction algorithm for disease dynamics based on mobile big data. *Information Sciences*, 379, 82–93. doi:10.1016/j.ins.2016.07.075

Chen, Y. and Yang, H., 2016. Sparse modeling and recursive prediction of space–time dynamics in stochastic sensor networks. *IEEE Transactions on Automation Science and Engineering*, 13 (1), 215–226. doi:10.1109/TASE.2015.2459068

Crandall, J.W., *et al.*, 2018. Cooperating with machines. *Nature Communications*, 9 (1), 233. doi:10.1038/s41467-017-02597-8

Deng, Z., *et al.*, 2015. A scalable and fast OPTICS for clustering trajectory big data. *Cluster Computing*, 18 (2), 549–562. doi:10.1007/s10586-014-0413-9

Fang, Z., *et al.*, 2017. Spatiotemporal model for assessing the stability of urban human convergence and divergence patterns. *International Journal of Geographical Information Science*, 31 (11), 2119–2141. doi:10.1080/13658816.2017.1346256

Frumkin, H., 2016. Urban sprawl and public health. *Public Health Reports*. doi:10.1093/phr/117.3.201

Fu, C., *et al.*, 2018. Identifying spatiotemporal urban activities through linguistic signatures. *Computers, Environment and Urban Systems*, 72, 25–37. doi:10.1016/j.compenvurbsys.2018.07.003

Fuzzi, S., *et al.*, 2015. Particulate matter, air quality and climate: lessons learned and future needs. *Atmospheric Chemistry and Physics*, 15 (14), 8217–8299.

Gao, J., *et al.*, 2018. On threshold-free error detection for industrial wireless sensor networks. *IEEE Transactions on Industrial Informatics*, 14 (5), 2199–2209. doi:10.1109/TII.9424

Garcia Alvarez, M., Morales, J., and Kraak, M. J., 2019. Integration and exploitation of sensor data in smart cities through event-driven applications. *Sensors*, 19 (6), 1372.

Gong, L., *et al.*, 2016. Inferring trip purposes and uncovering travel patterns from taxi trajectory data. *Cartography and Geographic Information Science*, 43 (2), 103–114. doi:10.1080/15230406.2015.1014424

Gong, P., *et al.*, 2012. Urbanisation and health in China. *The Lancet*, 379 (9818), 843–852. doi:10.1016/S0140-6736(11)61878-3

Gong, S., *et al.*, 2019. Extracting activity patterns from taxi trajectory data: a two-layer framework using spatio-temporal clustering, Bayesian probability and Monte Carlo simulation. *International Journal of Geographical Information Science*, 1–25. doi:10.1080/13658816.2019.1641715

Harris, N.L., *et al.*, 2017. Using spatial statistics to identify emerging hot spots of forest loss. *Environmental Research Letters*, 12 (2), 024012. doi:10.1088/1748-9326/aa5a2f

He, J., *et al.*, 2019a. Diverse visualization techniques and methods of moving-object-trajectory data: a review. *ISPRS International Journal of Geo-Information*, 8 (2), 63. doi:10.3390/ijgi8020063

He, Z., *et al.*, 2017. Mapping to cells: a simple method to extract traffic dynamics from probe vehicle data. *Computer-Aided Civil and Infrastructure Engineering*, 32 (3), 252–267. doi:10.1111/mice.2017.32.issue-3

He, Z., *et al.*, 2019b. Mining spatiotemporal association patterns from complex geographic phenomena. *International Journal of Geographical Information Science*, 1–26. doi:10.1080/13658816.2019.1566549

Hu, F., *et al.*, 2018. ClimateSpark: an in-memory distributed computing framework for big climate data analytics. *Computers & Geosciences*, 115, 154–166. doi:10.1016/j.cageo.2018.03.011

Iqbal, M.A., *et al.*, 2019. Amazon cloud computing platform EC2 and VANET simulations. *International Journal of Ad Hoc and Ubiquitous Computing*, 30 (3), 127–136. doi:10.1504/IJAHUC.2019.098472

Kan, Z., *et al.*, 2018. Fine-grained analysis on fuel-consumption and emission from vehicles trace. *Journal of Cleaner Production*, 203, 340–352. doi:10.1016/j.jclepro.2018.08.222

Karim, S., Soomro, T.R., and Burney, S.A., 2018. Spatiotemporal aspects of big data. *Applied Computer Systems*, 23 (2), 90–100. doi:10.2478/acss-2018-0012

Kim, Y.L., 2019. Data-driven approach to characterize urban vitality: how spatiotemporal context dynamically defines Seoul's nighttime. *International Journal of Geographical Information Science*, this issue. 1–22. doi:10.1080/13658816.2019.1694680

Kissling, W.D., *et al.*, 2018. Building essential biodiversity variables (EBV s) of species distribution and abundance at a global scale. *Biological Reviews*, 93 (1), 600–625. doi:10.1111/brv.12359

Kitchin, R., 2019. The timescape of smart cities. *Annals of the American Association of Geographers*, 109 (3), 775–790. doi:10.1080/24694452.2018.1497475

Kotevska, O., *et al*., 2017. Dynamic network model for smart city data-loss resilience case study: city-to-city network for crime analytics. *IEEE Access*, 5, 20524–20535. doi:10.1109/ACCESS.2017.2757841
Kraemer, M.U., *et al*., 2018. Inferences about spatiotemporal variation in dengue virus transmission are sensitive to assumptions about human mobility: a case study using geolocated tweets from Lahore, Pakistan. *EPJ Data Science*, 7 (1), 16. doi:10.1140/epjds/s13688-018-0144-x
Li, F., *et al*., 2018. Big enterprise registration data imputation: supporting spatiotemporal analysis of industries in china. *Computers. Environment and Urban Systems*, 70, 9–23.
Li, F., *et al*., 2018. Big enterprise registration data imputation: supporting spatiotemporal analysis of industries in China. *Computers, Environment and Urban Systems*, 70, 9–23. doi:10.1016/j.compenvurbsys.2018.01.010
Li, R., *et al*., 2017a. A replication strategy for a distributed high-speed caching system based on spatiotemporal access patterns of geospatial data. *Computers, Environment and Urban Systems*, 61, 163–171. doi:10.1016/j.compenvurbsys.2014.02.009
Li, Z., *et al*., 2017b. A high performance query analytical framework for supporting data-intensive climate studies. *Computers, Environment and Urban Systems*, 62, 210–221. doi:10.1016/j.compenvurbsys.2016.12.003
Li, Z., *et al*., 2019. SOVAS: a scalable online visual analytic system for big climate data analysis. *International Journal of Geographical Information Science*. doi:10.1080/13658816.2019.1605073
Liu, X., *et al*., 2015. Understanding urban China with open data. *Cities*, 47, 53–61. doi:10.1016/j.cities.2015.03.006
Liu, X., *et al*., 2018b. High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google earth engine platform. *Remote Sensing of Environment*, 209, 227–239. doi:10.1016/j.rse.2018.02.055
Liu, X.N., *et al*., 2019, March. Detection on abnormal usage of spectrum by electromagnetic data mining. *In*: *2019 IEEE 4th International Conference on Big Data Analytics (ICBDA)*. Suzhou, China: IEEE, 182–187.
Liu, Z., *et al*., 2018a. Mapping hourly dynamics of urban population using trajectories reconstructed from mobile phone records. *Transactions in GIS*, 22 (2), 494–513. doi:10.1111/tgis.2018.22.issue-2
Losser, T., Li, L., and Piltner, R., 2014, A spatiotemporal interpolation method using radial basis functions for geoBig Spatiotemporal Data. *In*: *2014 fifth international conference on computing for geospatial research and application*. Washington, DC: IEEE, 17–24.
Machado, K.L., *et al*., 2019. A data-centric approach for social and spatiotemporal sensing in smart cities. *IEEE Internet Computing*, 23 (1), 9–18. doi:10.1109/MIC.2018.2881517
Martín, Y., Li, Z., and Cutter, S.L., 2017. Leveraging twitter to gauge evacuation compliance: spatiotemporal analysis of hurricane matthew. *Plos One*, 12 (7), e0181701.
Martínez–Álvarez, F., and Morales–Esteban, A., 2019. Big data and natural disasters: new approaches for spatial and temporal massive data analysis. 129, 38–39.
McKinley, D.C., *et al*., 2017. Citizen science can improve conservation science, natural resource management, and environmental protection. *Biological Conservation*, 208, 15–28. doi:10.1016/j.biocon.2016.05.015
Moon, A., *et al*., 2017, September. Lossy compression on IoT big data by exploiting spatiotemporal correlation. In *2017 IEEE High Performance Extreme Computing Conference (HPEC)*. Waltham, MA: IEEE, 1–7.
Njuguna, C. and McSharry, P., 2017. Constructing spatiotemporal poverty indices from big data. *Journal of Business Research*, 70, 318–327. doi:10.1016/j.jbusres.2016.08.005
Pappalardo, L. and Simini, F., 2018. Data-driven generation of spatio-temporal routines in human mobility. *Data Mining and Knowledge Discovery*, 32 (3), 787–829. doi:10.1007/s10618-017-0548-4
Park, Y. and Kwan, M.P., 2017. Multi-contextual segregation and environmental justice research: toward fine-scale spatiotemporal approaches. *International Journal of Environmental Research and Public Health*, 14 (10), 1205.
Pavlyuk, D., 2019. Feature selection and extraction in spatiotemporal traffic forecasting: a systematic literature review. *European Transport Research Review*, 11 (1), 6. doi:10.1186/s12544-019-0345-9

Peters, D.P., *et al*., 2018. An integrated view of complex landscapes: a big data-model integration approach to transdisciplinary science. *BioScience*, 68 (9), 653–669. doi:10.1093/biosci/biy069

Rund, S.S., *et al*., 2019. Rescuing troves of hidden ecological data to tackle emerging mosquito-borne diseases. *Journal of the American Mosquito Control Association*, 35 (1), 75–83. doi:10.2987/18-6781.1

Samuel, A., *et al*., 2015. A framework for composition and enforcement of privacy-aware and context-driven authorization mechanism for multimedia big data. *IEEE Transactions on Multimedia*, 17 (9), 1484–1494. doi:10.1109/TMM.6046

Schimbinschi, F., *et al*., 2017. Topology-regularized universal vector autoregression for traffic forecasting in large urban areas. *Expert Systems with Applications*, 82, 301–316. doi:10.1016/j.eswa.2017.04.015

Shekhar, S., *et al*., 2015. Spatiotemporal data mining: a computational perspective. *ISPRS Int. J. Geo-Inf.*, 4, 2306–2338. doi:10.3390/ijgi4042306

Srivastava, S., *et al*., 2018. Fine-grained landuse characterization using ground-based pictures: a deep learning solution based on globally available data. *International Journal of Geographical Information Science*, 1–20. doi:10.1080/13658816.2018.1542698

Tang, L., *et al*., 2019. Detecting and evaluating urban clusters with Big Spatiotemporal Data. *Sensors*, 19 (3), 461. doi:10.3390/s19030461

Wang, S., Zhong, Y., and Wang, E., 2019. An integrated GIS platform architecture for Spatiotemporal Big Data. *Future Generation Computer Systems*, 94, 160–172. doi:10.1016/j.future.2018.10.034

Wei, J., *et al*., 2018. Spatiotemporal fusion of modis and landsat-7 reflectance images via compressed sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 55 (12), 7126–7139. doi:10.1109/TGRS.2017.2742529

Wei, S., *et al*., 2017. Exploring the potential of open big data from ticketing websites to characterize travel patterns within the Chinese high-speed rail system. *PloS One*, 12 (6), e0178023. doi:10.1371/journal.pone.0178023

Williams, B., *et al*., 2018. 2018 ESC/ESH guidelines for the management of arterial hypertension: the task force for the management of arterial hypertension of the European society of cardiology and the European society of hypertension. *Journal of Hypertension*, 36 (10), 1953–2041.

Xia, J., Yang, C., and Li, Q., 2018. Using spatiotemporal patterns to optimize earth observation big data access: novel approaches of indexing, service modeling and cloud computing. *Computers, Environment and Urban Systems*, 72, 191–203. doi:10.1016/j.compenvurbsys.2018.06.010

Xie, J., *et al*., 2019a. Similarity search of spatiotemporal scenario data for strategic air traffic management. *Journal of Aerospace Information Systems*, 187–202. doi:10.2514/1.I010692

Xie, Y., *et al*., 2019b. Modeling grassland ecosystem responses to coupled climate and socioeconomic influences in multi-spatial-and-temporal scales. *Journal of Environmental Informatics*, 33, 1.

Xu, J., *et al*., 2015a. Mining the situation: spatiotemporal traffic prediction with big data. *IEEE Journal of Selected Topics in Signal Processing*, 9 (4), 702–715. doi:10.1109/JSTSP.2015.2389196

Xu, Y., *et al*., 2015. Classifying dynamic textures via spatiotemporal fractal analysis. *Pattern Recognition*, 48 (10), 3239–3248.

Xu, Y., *et al*., 2015b. Classifying dynamic textures via spatiotemporal fractal analysis. *Pattern Recognition*, 48 (10), 3239–3248. doi:10.1016/j.patcog.2015.04.015

Xu, Z., Hu, C., and Mei, L., 2016. Video structured description technology based intelligence analysis of surveillance videos for public security applications. *Multimedia Tools and Applications*, 75 (19), 12155–12172. doi:10.1007/s11042-015-3112-5

Yan, J., *et al*., 2017. Industrial big data in an industry 4.0 environment: challenges, schemes, and applications for predictive maintenance. *IEEE Access*, 5, 23484–23491. doi:10.1109/ACCESS.2017.2765544

Yang, C., *et al*., 2011. Using spatial principles to optimize distributed computing for enabling the physical science discoveries. *Proceedings of the National Academy of Sciences*, 108 (14), 5498–5503. doi:10.1073/pnas.0909315108

Yang, C., *et al*., 2017. Utilizing cloud computing to address big geospatial data challenges. *Computers, Environment and Urban Systems*, 61, 120–128. doi:10.1016/j.compenvurbsys.2016.10.010

Yang, C., *et al*., 2019. Big Earth Data Analytics: a survey. *Big Earth Data*, 3 (2), 83–107. doi:10.1080/20964471.2019.1611175

Yang, J.A., *et al*., 2016. Social media analytics and research testbed (SMART): exploring spatiotemporal patterns of human dynamics with geo-targeted social media messages. *Big Data & Society*, 3 (1), 2053951716652914. doi:10.1177/2053951716652914

Zhang, P., *et al*., 2016. Deformable object tracking with spatiotemporal segmentation in big vision surveillance. *Neurocomputing*, 204, 87–96. doi:10.1016/j.neucom.2015.07.149

Zhang, T., *et al*., 2019. Identifying primary public transit corridors using multi-source big transit data. *International Journal of Geographical Information Science*. doi:10.1080/13658816.2018.1554812

Zhang, X., *et al*., 2018. Geospatial sensor web: a cyber-physical infrastructure for geoscience research and application. *Earth-science Reviews*, 185, 684–703. doi:10.1016/j.earscirev.2018.07.006

Zhao, P., *et al*., 2018. An empirical study on the intra-urban goods movement patterns using logistics big data. *International Journal of Geographical Information Science*, 1–28. doi:10.1080/13658816.2018.1520236

Zhu, Y., *et al*., 2016. Using big data to study resilience of taxi and subway trips for hurricanes Sandy and Irene. *Transportation Research Record*, 2599 (1), 70–80. doi:10.3141/2599-09

Chaowei Yang
NSF Spatiotemporal Innovation Center and Department of Geography and GeoInformation Science, George Mason University, Fairfax, VA, USA
cyang3@gmu.edu http://orcid.org/0000-0001-7768-4066

Keith Clarke
Department of Geography, University of California, Santa Barbara, CA, USA
http://orcid.org/0000-0001-5805-6056

Shashi Shekhar
Department of Computer Science, University of Minnesota, Minneapolis, MN, USA

C. Vincent Tao
Wuhan University and Z Ventures Group, Shanghai, PR China