Research and Application Exploration for Spatiotemporal Data Integrated Theory

Yuan Chao1*, Zeng Tianye2, Cheng Yuxiang1, Li Lin1

1 Chongqing Geomatics and Remote Sensing Center, Chongqing 401147, CHINA
2 Oberlin College, OH, USA

*Corresponding author’s e-mail: cqdlkj@126.com

Abstract. Spatiotemporal data is essential for big data, closely related to government, enterprise and people. The Chinese digital economy development strategy requires more efficient usage of spatiotemporal data. However, current researchers paid the more attention to the concept and technology about spatiotemporal database, model, sharing, visual platform, lacking research to the spatiotemporal data integrated theory. This article attempted to propose the spatiotemporal data integrated theory, and structure three-level spatiotemporal data system. It solved systematically spatiotemporal data problem about classification, calculation and application, developed series smart software system, formed a spatiotemporal big data factory in Chongqing China.

1. Introduction
Spatiotemporal data means spatial data with time series, and it is the most important big data[1]. In our daily life, spatial data with time series is everywhere, and 80% of the data produced by human's life are related to spatial position[2]. In 2011, Manyika from McKinsey Global Institution published the report "big data: the next frontier for innovation, competition, and productivity." In the report, he claims that health care, retail, public domain, manufacture, and personal location comprise the popular big data stream at that time, and all of those data have great spatiotemporal characteristics.

With the appearance of big data, cloud calculation, networking, and AI, the number of spatiotemporal data increases exponentially. Location information, movement trace, natural change, and human activities, which reflect economic, social, human, ecological information, become spatiotemporal data of sensed, saved, and analyzed. The sensor we can get spatiotemporal data varies from professional sensors to networking ubiquitous unprofessional sensors. Aerial and satellite image, city real image, city video data, and moving trace are examples of such sensors.

How to process and analyze spatiotemporal data efficiently is technical focus in big data time. With the developed big data analysis technology and the learning ability of AI, it is not impossible to uncover the massive knowledge behind the spatiotemporal data. Besides, multi-source spatiotemporal data and related areas create numerous spatiotemporal data derivative products such as cloud center, spatiotemporal model and database, spatiotemporal data analysis method, spatiotemporal data real time calculation, spatiotemporal data safe system and spatiotemporal data application case. These derivative products bring favorable spatiotemporal data primary environment, which can give solutions for various professions about spatiotemporal data real time connection, integrated management, and high effective calculation. However, most of spatiotemporal data related solutions are scattered and conceptual. Most of these are about concepts and cloud platform development.
technology [3-9], but limitedly use spatiotemporal data theory system and spatiotemporal data integrated theory. The writer and research group use their years' experience of spatiotemporal data management, and sharing application proposed the three-level spatiotemporal data system developed, and built the spatiotemporal data factory in Chongqing China. It became the spatiotemporal pivot of smart Chongqing.

2. Spatiotemporal Data Integrated Theory Research

The spatiotemporal data is massive, in different types, difficult to process, and update frequently. All those features make high requirements for the technical workers' profession, the efficiency of calculation storage equipment, and the decrease of the application cost. How can spatiotemporal data classify intelligently, calculated fast, applied iteratively? This research proposes the theory of classification code, computing grid and management unit integration progressively.

2.1. Spatiotemporal Data Classification Code Theory

Classification is the most primary way for humans to know the world. In order to understand enormous data, we should sort out data and do classify identification. In this research, we use geographical entity as the smallest unit and classify spatiotemporal data based on its attributes. In Chongqing city we divide urban spatiotemporal data into 5 categories: geography data, surface data, planning data, economic data, and city operating data. In them, geography data includes land utilization, building, facility, underground structure, and geographical unit. Surface data comprises data in land condition, house, infrastructure, public management and service facilities, resources and environment. Planning data include national planning, economic and society development planning, urban and rural planning, land planning, and environmental protection planning. Planning data show the urban information for the future. Economic data conclude national economy account, population, employment, assets investment, finance, people’s living, and price fluctuation. It express urban economic activities. City operating data is mainly about population, corporation, traffic, resource, project, and emergence. It conveys the dynamic spatiotemporal information in city.

In this research, the spatiotemporal data’s category code consists of geographical entity code and attribute code. The geographical entity code is the category code of geographical entity, and the attribute code means the category code of geographic entity’s natural, social, and dynamic feature. With the combination of geographical entity code and attribute code, we can actualize the unified classification of all spatiotemporal data.

In the figure below, the geography data, surface data, planning data, economic data, and city operating data are names A, B, C, D, and E correspondingly. Type A, geographical entity code uses 9 digits: 1 digit for the category, 2 digits for the first subclass, 2 digits for the second subclass, 2 digits for the third subclass, and 2 digits for the fourth subclass. The code structure is the figure below.

![Figure 1. Code Structure of Geographic Entity Category](image)

Take “intercity road” as an instance of Type A code. The classification code for “intercity road” is “A03010201”: “A” stands for geographic data, “03” means the first subclass “infrastructure,” “01”...
means the second subclass “traffic and facility,” “02” means the third subclass “traffic way and facility,” and “01” means the fourth subclass “intercity road.” The detail are in the following figure:

| 1st-subclass | 2nd-subclass | 3rd-subclass | 4th-subclass | Category code | description |
|--------------|--------------|--------------|--------------|---------------|-------------|
| Facility (03) |              |              |              | A030000000    | it is material engineering facilities providing public services for social production and residents’ life |
| Traffic facilities (01) |              |              |              | A030100000    | General term for railway facilities, highway facilities, water transport facilities and aviation facilities |
| Railway facilities (01) |              |              |              | A030101000    | Tracks and ancillary facilities for trains |
| Highway facilities (02) |              |              |              | A030102000    | Including all levels of roads for vehicles driving and additional facilities such as protection, drainage, maintenance, management, service, safety, ferry transportation, monitoring, communication, toll collection, etc. |
| Intercity highway (01) |              |              |              | A030102010    | A road Connects two cities |

Figure 2. Example for Geographic Entity Category Code

Type B data, surface data, comprises 13 digits code about the natural and social attributes of a geographical entity. Combine the geographical entity code with surface category code, which means append Type B data to Type A data. The geographical entity would gain the attributes of the surface status.

For example, appending the “high-speed railway” attribute to “railway and facility,” the code would be “A03010100 B040101010101”. “A03010100” stands for geographical entity “railway and facility”. “B040101010101” presents the surface attribute code “high-speed railway”. See the chart below:

| 1st-subclass | 2nd-subclass | 3rd-subclass | 4th-subclass | Category code | description |
|--------------|--------------|--------------|--------------|---------------|-------------|
| Facility (03) |              |              |              | A030000000    | it is material engineering facilities providing public services for social production and residents’ life |
| Traffic facilities (01) |              |              |              | A030100000    | General term for railway facilities, highway facilities, water transport facilities and aviation facilities |
| Railway facilities (01) |              |              |              | A030101000    | Tracks and ancillary facilities for trains |
| Highway facilities (02) |              |              |              | A030102000    | Including all levels of roads for vehicles driving and additional facilities such as protection, drainage, maintenance, management, service, safety, ferry transportation, monitoring, communication, toll collection, etc. |
| Intercity highway (01) |              |              |              | A030102010    | A road Connects two cities |

Figure 3. Example for Category Code of The Surface Data
Rather than regurgitating, the code rule of planning data, economic data, and city operating data (Type C, D, and E) are similar to the surface data (Type B).

2.2. Spatiotemporal Data Computing Grid Theory
In the traditional survey, the primary geographic grids are the carrier of spatiotemporal data. After processing spatiotemporal data, grids are used to be the unit to store and manage the data. By the relative size to the grid center show the location [10-13]. This research bottoms on the classification of spatiotemporal data, with the geographic entity’s location and spatial area information, we built computing grid. It forms a new spatiotemporal index and can quick search massive spatiotemporal data, converts more complex spacial geometry calculating to simple coding computing. It can achieve higher efficiency of spatiotemporal analysis.

This research uses the GeoHash coding to convert spatial coordinates to the global grid[14,15], then extract binary grid code, and combine longitude and latitude binary code. After that convert to text code by base32 coding, finally the spatiotemporal text code are formed. The set of encoding saved numerous floating-point computing, enhanced the efficiency of spatial calculating, meanwhile protected the real coordinate information. In Figure 4, it subdivides longitude 106.503, 0 means on the left side of the subdivision point, and 1 means on the right side of the subdivision point. The longitude binary code we get it 11001, the more level we have, the higher the accuracy.

| Level | 106.503 | Longitude range | Interval (0) | Interval (1) |
|-------|---------|-----------------|--------------|--------------|
| 1     | 1       | (-180, 180)     | (-180,0)     | (0,180)      |
| 2     | 1       | (0,180)         | (0,90)       | (90,180)     |
| 3     | 0       | (90,180)        | (90,135)     | (135,180)    |
| 4     | 0       | (90,135)        | (90,112.5)   | (112.5,135)  |
| 5     | 1       | (90,112.5)      | (90,101.25)  | (101.25,112.5)|

Figure 4. Example for Longitude Binary Coding

In the same way, the binary coding of latitude is 10101, Then if we reorganize them, make the even digits present the longitude and odd digits for latitude, we can get the grid binary code 11011 00011. Then, the binary code is combined from left to right according to 5 bits, and the minimum is 0, and the maximum is 31. The 5-bit combination is converted to decimal code, corresponding to (27, 3); finally, the base32 coding is used to convert the decimal into string expression (V3). So these many coding conversion can protect the real coordinate information to a certain extent.

| Decimal  | 0   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  | 14  | 15  |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Base32   | 0   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | b   | c   | d   | e   | f   | g   |
| Decimal  | 16  | 17  | 18  | 19  | 20  | 21  | 22  | 23  | 24  | 25  | 26  | 27  | 28  | 29  | 30  | 31  |
| Base32   | h   | j   | k   | m   | n   | p   | q   | r   | s   | t   | u   | v   | w   | x   | y   | z   |

Figure 5. Example for Base32 Coding

2.3. Spatiotemporal Data Management Unit Integration Theory
The reason that spatiotemporal data research becomes hotspot in the big data era is its immeasurable value. To satisfy different application requirement, this study use management unit as the carrier, by management unit code to integrate spatiotemporal data and service the management job in government department. Chinese general management unit includes administrative divisions, population census units, economic census units, community grid, forestry compartment, and traffic zone[16,17], These are the smallest work unit serving for multiple management.
The rule of managing units coding vary for different applications. For instance, The Chinese administrative division has already been inherent coding. The administrative division code comprised of a 12 digits code, divides the province, county, subdistrict, community four levels, as the following graph.

![Figure 6. Structure of Chinese Administrative Division Code](image)

The community grid has played an essential role in 2020 prevention and control of COVID-19. Its code structure is appending 3 digits code after the administrative division code, as shown in the figure below.

![Figure 7. The Chinese Society Management Grid Code Structure](image)

We mark spatiotemporal data in the management area of their corresponding unit code when facing different applications based on spatiotemporal data grid code. This way, we get spatiotemporal dataset with management marks, so we can focus on spatiotemporal management applications, leading to significantly improve efficiency of spatiotemporal data inputting, analysing, and outputting.

3. Three-Level Spatiotemporal Data System Construct
The spatiotemporal data integrated theory in this research solves what is, where is, and how to use spatiotemporal objects through building spatiotemporal data classification code, grid code, and unit code. In the last two years, we have applied this theory to construct China Chongqing city spatiotemporal data system. We have developed a set of software tools, including spatialization converting, auto coding, quality checking, data integrating, spatial map analysis. Those tools are used to build China Chongqing spatiotemporal big data factory and construct three-level spatiotemporal data products.
Figure 8. Three-level Architecture of Spatiotemporal Data in China Chongqing

The first level spatiotemporal data. First of all, the spatiotemporal objects in Chongqing are sorted out in terms of longitude and latitude. If the spatiotemporal objects are in text format, they should be spatialized, and the location description information should be transformed into longitude and latitude coordinates. Then, according to the spatiotemporal data classification code rules of this study, the spatiotemporal objects are classified and identified, which are divided into five categories: geography data, surface data, planning data, economic data, and city operating data, so that each spatiotemporal object has a unique classification code.

The second level spatiotemporal data. The classified spatiotemporal data is converted to geographic grid, and the grid code of the location of spatio temporal objects is marked. A secure and efficient spatiotemporal index is constructed to realize fast search access and spatial analysis. At the same time, the spatiotemporal data is mapped to the global grid system. After several code converting, the text coding is formed, and the real coordinates are protected to facilitate spatio temporal data sharing applications.

The third level spatiotemporal data. Spatiotemporal data is widely used in many fields, but its efficiency and value are limited now. Based on the theory of spatiotemporal data management unit integrating, this study integrates spatiotemporal data by the smallest work unit, identifies the spatiotemporal data within the same management unit with the same unit code, to build the third level spatiotemporal data. Through the three-level spatiotemporal data, users can make decision-making layout, implement evaluation and supervision under unified management standards, and achieve accurate management in spatiotemporal area according to different management demand.[18,19]

4. Three-Level Spatiotemporal Data Products Applicable

The exploration of spatiotemporal data integrated theory in this study has been fully demonstrated in Chongqing China, forming three-level spatiotemporal data system product of the whole city. Using the map as the "view" representation of the real world[20], it has attempted new application different from traditional geographic information service.

4.1. Chongqing COVID-19 Epidemic Map Application

In the global fighting of COVID-19 in 2020, there appeared multiple versions of epidemic maps. These maps use to show the epidemic situation in maps with big-data analysis and allow the public to know about the latest prevention and control progress. On 2020 January 24, the Chongqing city government started first-level emergency response. On February 13th, our research group brought out the “Chongqing City Epidemic Map” on WeChat. This app pushed real-time the spatiotemporal epidemic information and got over 7 million hits quickly with full acceptance. We could bring out the
map in such a short time that benefited from ability of three-level spatiotemporal data products and Chongqing city spatiotemporal big-data factory.

Specifically, we used the first level spatiotemporal data tool to convert the address and activity locations of confirmed cases from text into spatiotemporal data, getting the patient’s activity track. After this, we integrated the third level spatiotemporal data product to connect the patient’s location with management unit of residential district. In this way, we successfully refined the distribution of the confirmed cases in the whole city and provided citizens a convenient means to know details. Besides, we connect confirmed cases with administrative division management units, divide the epidemic in the city into different regions and classes. It services effectively epidemic grid management.

4.2. Spatiotemporal Brain Application for Industry

Like the promotion of digital economy strategy and Structural adjustment and layout of strategic emerging industries in China, there are imperious demands for accurate implementation and multidimensional visualization of industrial planning. Modern enterprises' development continues to expand to the whole industry chain and high technology. A single production function no longer meets the needs of enterprise development but needs to integrate production, research and development, display, marketing, service, and other compound functional space. In order to realize the intelligent guidance and precise matching of various industrial elements and resources in Chongqing China, this study proposes the "spatiotemporal brain for industry", which is mainly based on the three-level spatiotemporal data system, integrates the industrial chain analysis results and multidimensional industrial information. This study shaped spatiotemporal state for the industry in Chongqing, and analyzed the advantages and disadvantages of industry developing on resources, enterprises, parks, and industries.

This study draws the industrial spatiotemporal map from multiple perspectives of industry, region, and enterprise. It forms knowledge achievements such as key enterprise maps, industrial chain spatiotemporal map, regional industry map, precise investment map, industrial development park map. The spatiotemporal brain will support the application of modern industrial layout, accurate investment attraction, investment project management, and industrial policy implementation.

5. Conclusion

Human activity is essentially a kind of spatiotemporal behavior. The occurrence, development, and evolution of everything and phenomenon in the world are carried out in a particular time and space. From the birth of the world's first map (the first map of Babylon 4500 years ago) to the rapid development of modern science and technology, human beings have been exploring the natural laws of spatiotemporal data and its universal value. With the development of science and technology for thousands of years, from "experimental science" to today's "data science," what are the hidden regular pattern of spatiotemporal big data? Many experts and scholars have spent their whole life researching and exploring, but unfortunately, the effect is not apparent. Starting from another research idea, "if there is no rule, then establish rules," this paper explores and forms a set of rules for the integration, processing, and service of spatiotemporal data, which makes multi-source and heterogeneous spatiotemporal big data obtain a clue. Under the unified rules, unified classification, rapid calculation, and flexible application, these rules focus on the knowledge value of spatiotemporal data. For a long time, spatiotemporal data has been confined to the positioning of the primary carrier and limited to map basic attribute. It is time to consider the promotion of spatiotemporal data from primary products to secondary and tertiary products, to realize the maximum release of spatiotemporal value. This paper studies the theory and practice of Chongqing city in China and has made a decisive step. I hope our work can get more attention and support, learn from other, and enjoy value of spatiotemporal data.

Acknowledgments

This paper is supported by the Special Fund for Technology Innovation and Application Development of Chongqing, China (Grant No. cstc2019jsx-gksbX0042)
References

[1] LI, D. (2016) Towards geo-spatial information science in big data era[J]. Acta Geodaetica et Cartographica Sinica, 45:379-384.

[2] Xu, G.H. (1999) Pay much attention to the digital earth by the social[J]. Science News Weekly,1: 7-8.

[3] LI, D., MA, J., SHAO, Z.F. (2015) On Spatio-temporal Big Data and Its Application[J]. Satellite Application, 9:7-11.

[4] Wang, J.Y., Wu, F., Guo, J.Z., et al. (2017) Challenges and opportunities of spatio-temporal big data[J]. Science of Surveying and Mapping, 42:1-7.

[5] Bian, F.L., Du, J.Y., Meng, X.L. (2016) Requirements, Applications and Challenges of Spatio-Temporal Big Data Processing[J]. Journal of Geomatics, 4:1-4.

[6] Xiao, J.H., Wang, H.Z., Peng, Q.S., et al. (2016) Research on the Construction of Cloud Platform for the Spatio-temporal Big Data Management and Application[J]. Bulletin of Surveying and Mapping, 4:38-42.

[7] LIN Hui, YOU Lan, et al. Prospect of geographic knowledge engineering in the era of spatiotemporal big data [J]. Journal of Wuhan University, 2018,43:2205-2211.

[8] Wang, J.Y., (2017) Cartography in the era of spatiotemporal big data [J]. Acta Geodaetica et Cartographica Sinica, 46:1226-1237.

[9] OGC, (2003) Overview of OGC’s Interoperability Program.

[10] Dutton, G. (2000). Universal Geospatial Data Exchange via Global Hierarchical Coordinates. International Conference on Discrete Global Grids, Santa Barbara, California, USA. 3.

[11] White D., Kimerling, J., etc. (1998) Comparing area and shape distortion on polyhedral-based recursive partitions of the sphere. Int J Geographical Information Science, 12: 805- 827.

[12] Li, D, Xiao, Z.F., et al. (2006) Research on multi grid partition and coding of spatial information [J]. Acta Geodaetica et Cartographica Sinica, 31:52-56.

[13] Hu, L.J., Cai, J., et al. (2018) Spatial data fusion method based on spatiotemporal geographic grid [J]. Journal of Survey and Geospatial Information, 41:4-7.

[14] Goodchild, M. (2001) A geographer looks at spatial information theory. In: Spatial Information Theory: foundation of geographic information science. Springer-Verlag. Berlin, Heidelberg, New York. PP:1-13.

[15] Gerstner, T. (2002) Multiresolution Visualization and Compression of Global Topographic Data. Geoinformatic, http://wissrech.iam.uni-bonn.de/research/pub/gerstner/globe.pdf.

[16] Li, X.D., Yang, X.G., et al. (2009) Research on traffic zone division method of urban road network [J]. Journal of Computer Engineering and Application, 45:19-22.

[17] Zhu, J., Li S.J., et al. (2018) Exploration and suggestions on the innovation of industrial development unit planning and management -- Taking Hangzhou as an example [J]. Journal of Zhejiang Construction, 35:1-5.

[18] Wang W, Hu C, Chen N, et al. (2016) Spatio-Temporal Risk Assessment Process Modeling for Urban Hazard Events in Sensor Web Environment[J]. International Journal of Geo-Information, 5:203.

[19] Wei W, Hu C B, Chen N C, et al. (2015) Spatio-Temporal Enabled Urban Decision-Making Process Modeling and Visualization Under the Cyber-Physical Environment[J]. Science China Information Sciences , 58(10):1-17.

[20] Goodchild, M. (2000) Discrete Global Grids for Digital Earth. In: International Conference on Discrete Global Grids. Santa Barbara, California, USA.pp. 3.