

CONSTRUCTION METHOD OF “CELL-CUBE” SPATIO-TEMPORAL DATA MODEL FOR BIG DATA

Yanling.Lu^{1,2*}, Caiwei.Liu², Jingwen.Li^{1,2}, Jianwu.Jiang²

¹ Guangxi Key Laboratory of Spatial Information and Geomatics, Guilin University of Technology, Guilin, China-
Luy12014@glut.edu.cn

² College of Geomatics and Geoinformation, Guilin University of Technology, Guilin, China- 358498163@qq.com

Commission VI, WG VI/4

KEY WORDS: Big Data, Multi-Feature, Data Organization, Data Storage, Spatio-temporal Data Model

ABSTRACT:

In recent years, with high accuracy, high frequency, considerable coverage of remote sensing images, map tiles, video surveillance, web crawlers, social networking platforms and other types of spatiotemporal data have exploded in geometric progression. Human society has come into the era of big data in time and space. In view of the characteristics of multi-attribute, multi-dimensional, multi-source and heterogeneous spatiotemporal big data, how to make use of the emerging information technology means, combined with the geographic information data analysis means, the rapid mining and utilization of spatiotemporal big data has become a key problem to be solved. Built on the background of spatiotemporal big data and the process of geospatial cognition, this paper proposes a "cell-cube" spatiotemporal object data model. This paper constructs a model system of geo-spatiotemporal big data from the aspects of data organization, data storage and data partition, and abstracts the geo-space into an infinite number of geo-cells, and the adjacent geo-cells gather around the core cells to form geographical clusters, and the geographical clusters with similar attributes are clustered into geographical blocks. At the level of data organization, the spatial and temporal characteristics of structured data and unstructured data are considered as organizational dimensions, and a multi-factor extended cube data model is proposed. In the aspect of data storage, the organization model is further abstracted into the cell-cube structure of distributed data warehouse, and then the spatiotemporal data is stored uniformly. At the level of data segmentation, the mathematical table and space calculation method of multi-feature extended cube are proposed, and the geographical cell data division model based on connection is established. It solves the organization and management problem of spatiotemporal big data, provides a more complete data organization framework and solution for the application of geo-spatiotemporal big data, and promotes the development of deep mining of spatiotemporal extensive data in GIS. And to achieve space-time big data in the geographical space microscopic and the macroscopic unification cognition.

1. INTRODUCTION

At present, all walks of life around the world have entered the digital era. The rapid development of Smart City and cloud computing, as well as artificial intelligence and remote sensing technologies has made the amount of geographical data continuously expand. The space-time problems of diversity, access and consistency of space-time data need to be reconsidered. Based on remote sensing and cloud computing technology, Li Deren discusses complex feature mining of spatio-temporal big data. Based on the traditional characteristics of time, space and attribute, the multi-source, multi-structure, multi-scale and dynamic characteristics of spatiotemporal data are put forward. The whole spatial information system proposed by Hua Yixin is expanded to the analysis of big data space, and carries out the abstract management of various kinds of geographical spatial entities in a complex and dynamic way, and describes geographical phenomena or dynamic changes of

entities through the big data of space and time to achieve the real world space-time big data for a more scientific and practical expression.

Thus, the premise of spatio-temporal big data mining is how to express. It is difficult to organize and manage big data, but there are still many limitations in the cognitive style and organizational management of big data. Therefore, how to organize spatio-temporal big data effectively, quickly mine information and realize efficient management of spatio-temporal big data is a key scientific problem to be solved urgently. Based on the characteristics of spatio-temporal big data, this paper constructs a spatio-temporal big data model of the bionics structure of geospatial cells. The clustering of similar attributes provides a new idea for the organization and management of spatio-temporal data, and achieves the unity of micro and macro cognition and data organization of spatio-temporal data.

* Corresponding author

2. PRINCIPLE OF CELL BIOMIMETIC STRUCTURE

2.1 Definition of cell biomimetic structure

In traditional geospatial cognition, the expression of time is regarded as a dimension attribute or as a conditional context. Under the background of spatio-temporal big data, the cognitive model of the whole spatial GIS appears. Therefore, geospatial cognition has developed from three-dimensional to multi-dimensional, multi-volume, multi-source heterogeneous and has a close relationship with tension. Therefore, in this paper, the concept of the cell is used to express the spatio-temporal data of geographical entities based on the bionics structure of cells it has the characteristics of macro-continuity and micro-discreteness, which provide a new method for the space-time data management of hyper volume. The principle is: The Geographical Cell is the smallest unit in the cognitive space. Each geographical cell exists independently, and the geographical cells are related to each other. Among them, different geographical cells describe geographical entities in different time and space, and the time-space relationship based on cell bionics structure is expressed as a formula (1).

$$C_i = U(A, S(X, Y, Z), T, I) \quad (1)$$

In the formula(1), C_i represents the i geographical cell. The information stored in the cell includes attribute information, spatial information, time and time-varying event information of cell changes. Attribute information A is the basic attribute of geographical cells, such as elevation, slope, aspect and type of land. Spatial information mainly includes the spatial position and shape of geographical cells the time-integrated geographical cell can describe whether other information of geographical cell is different in different time by time series T; Time-varying event information I specifically describes the time when such information as geographic cell attributes change and the detailed description of events.

Each cell is an independent set of spatial position, attribute and time, and its cell bionic structure includes the following aspects:

(1) Time Series: a sequence $T = \{t_0, t_1, t_2, \dots, t_n\}$ consisting of a set of different time nodes in a geographical cell, where t_0 is regarded as an initial time or a reference time set according to different circumstances, and the time interval between two adjacent nodes is equal and is denoted as ΔT .

(2) Spatial Features: describe the specific location of geographical cells in space, in the coordinate system x-y-z, the spatial features of geographical cells in different time coordinate sequence as $S_i = (X_i, Y_i, Z_i), i \in [1, n]$.

(3) Attribute Table: Each geographical cell has the most basic attribute information of geographical entities, such as elevation, Slope, category, etc..

$$C_a = q_{ij} = (q_{i1} \quad q_{i2} \quad \dots \quad q_{im}) \quad (2)$$

(4) Spatial Morphology: the spatial shape of geographic cells after aggregation and combination, which can describe the shape of the ground object after generalization.

(5) Space-time Relationship: the relationship between geographical cells is not only the traditional spatial topological relationship, but also the time topological relationship. According to the first theorem of geography, for many geographical cells, the closer the spatial distance, the higher the similarity of their attributes. Similarly, with the same geographical cell as an object, the closer the time distance, the higher the similarity of their attributes.

(6) Spatio-temporal Events: Geographic cells are grouped together to form geographic clusters to express geographic entities, and events that occur in the real world are recorded as text within the cells, which can be represented by $I = \{I_0, I_1, I_2, \dots, I_n\}$.

The space-time events include the corresponding cell combination, the type, the cause, the beginning and the end time, the process, the periodicity, the space, the result. To change a field to a building site, geographical cells record changes in properties, locations, boundaries, and so on in the area at a time, emphasizing changes in time and space.

(7) Cognitive Process: The process and manner of geographical cell in geographical space cognition.

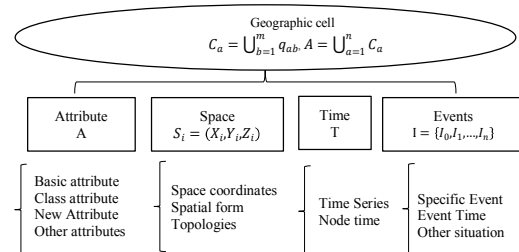


Fig. 1 Framework of geographical cell storage information

2.2 The relationship between cell biomimetic structure

There are a large number of independent geographic cells in the biomimetic structure of cells, which can express a kind of geographic entity by constructing the connection among the geographic cells with similar parameters but independent. The relationship between the bionics structure of cells is constructed through the similarity of the spatial-temporal information and the attribute information, and by establishing this relationship, the microscopical continuity of geographical cells is produced, and the relationship between geographical entities in the real world can be described in detail contribute to the cognitive process of geographical space. The organization and management of spatio-temporal big data can be improved by using the micro-to macro-characteristics of cell bionic structure.

The adjacent geographical cells have similar geographical attributes, and the geographical cells with similar attributes are clustered to form geographical clusters. In the cell bionic structure, the attributes of geographical clusters are stored by matrix A and matrix A describes in detail a geographical cell with N M attributes as follows:

$$A = (q_{ij})_{n \times m} = \begin{bmatrix} q_{11} & q_{12} & \dots & q_{1m} \\ q_{21} & q_{22} & \dots & q_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ q_{n1} & q_{n2} & \dots & q_{nm} \end{bmatrix} \quad (3)$$

In the formula(3), it can extract the attribute $C_a = q_{ij} = (q_{i1} \quad q_{i2} \quad \dots \quad q_{im})$ of a single geographical cell in a geographical cluster, which is denoted as set form $C_a = U_{b=1}^m q_{ab}, A = U_{a=1}^n C_a$. Because the units corresponding to the attributes in geographical cells may be different, in order to achieve more efficient and efficient computation, it is necessary to unify the units of each attribute value, that is, attribute normalization:

$$b_{ij} = \frac{p_{ij} - \min_j(p_{ij})}{\max_j(p_{ij}) - \min_j(p_{ij})} \quad (4)$$

The set of partial geographical cells in dataset A is recorded as:

$$g_a = \sum_{i=1}^n C_i, h \in [1, n] \quad (5)$$

In fact, the geographical cells in geographical clusters are in dynamic change. The increase and decrease of geographical cells directly affect the spatial form of geographical clusters and the attribute matrix of geographical clusters. Representing an existing geographic cluster as $G_i = U_{a=1}^i g_a$. In a time series, the new geographical cells cluster to the existing geographical cluster G_i . The geographic clusters change when $G_{i+1} = G_i \cup g_{a+1}$, a dynamic change that reflects the scalability of geographic clusters in the biomimetic structure of a cell, as shown in figure 2.

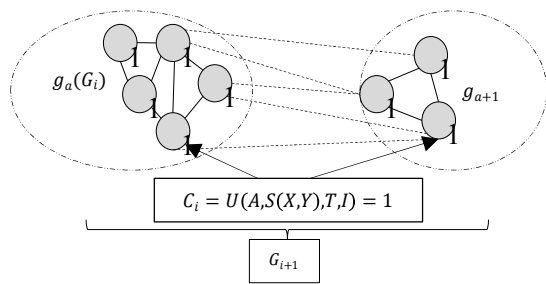


Fig. 2 Dynamic changes of geographical clusters

In this space-time sequence, the attribute information also changes dynamically, and the attribute information of geographical cluster is managed by the extended matrix.

$$\begin{matrix} G_{i-1} + g_a = G_i \\ \begin{bmatrix} q_{11} & q_{12} & \dots & p_{1m} \\ q_{21} & q_{22} & \dots & p_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ q_{n1} & q_{n2} & \dots & p_{nm} \end{bmatrix} + \begin{bmatrix} q_{1k+1} & q_{1k+2} & \dots & q_{1k+m} \\ q_{2k+1} & q_{2k+2} & \dots & q_{2k+m} \\ \vdots & \vdots & \vdots & \vdots \\ q_{nk+1} & q_{nk+2} & \dots & q_{nk+m} \end{bmatrix} = \\ \begin{bmatrix} q_{11} & q_{12} & \dots & q_{1m} & \dots & q_{1k+m} \\ q_{21} & q_{22} & \dots & q_{2m} & \dots & q_{2k+m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ q_{n1} & q_{n2} & \dots & q_{nm} & \dots & q_{nk+m} \end{bmatrix} \end{matrix} \quad (6)$$

3. GEO-SPATIAL COGNITION BASED ON CELL BIOMIMETIC STRUCTURE AGGREGATION

3.1 The aggregation process of cellular biomimetic structure

In order to facilitate the management of scattered geographical cells, the cube formed after clustering is the basic unit in the management of spatio-temporal big data by using the clustering feature of cell bionic structure, which reduces the management pressure of a large number of dense data geographical cells with the same properties are managed uniformly by clustering them into cubes. In addition, the management of the cube also facilitates the rapid division of new spatio-temporal data into suitable cubes and the establishment of links with other geographical cells. In the current spatial data clustering algorithm, density algorithm can mine and analyze arbitrary shape clusters. DBSCAN (Density-Based Spatial Clustering of Applications with Noise) Algorithm and OPTICS Algorithm are representative of the density algorithm. In this paper, based on DBSCAN Algorithm, from the point of view of geographical cells, the space-time domain is constructed to expand the definition of density clustering algorithm.

Taking core geographical cells with different attributes as the center, the neighborhood radius of core cells was set to Eps, and the height of core cells was time interval ΔT . The density of geographical cells contained in the neighborhood of core geographical cells can reach each other, that is to say, the density of geographical cells can reach by using core geographical cells as a bridge. Density clustering of geographical cells:

(1) Set the cell neighborhood radius Eps and equal interval time to determine whether the number of geographical cells in the cell neighborhood is less than the number of MinPts. Location of high density of geographical cells in space-time by setting threshold Density.

$$\text{Density} = \frac{\text{MinPts}}{\pi \cdot \text{Eps}^2 \cdot \Delta T} \quad (7)$$

(2) Traversing the data set G_i after attribute initial clustering. If the number of geographical cells is less than MinPts, the geographical cells are boundary cells and continue to query and search the remaining cells of set G_i . If it is larger than MinPts, the core cell C_k in the initial cluster can be found. The minimum distance from the core cell to other cells in its neighborhood is expressed as:

$$\text{core} - \text{dist} = \begin{cases} \text{Undefined, if } \text{Card}(N_\varepsilon(C_k)) < \text{MinPts} \\ \text{MinPts} - \text{dist}(C_k), \text{ Other situation} \end{cases} \quad (8)$$

(3) To solve the geographical cell with the core cell C_k density up to, so that it produces a new cell class W_i to represent the geographical entity composed of geographical cells. $N_\varepsilon(C_k)$ is the ε neighborhood of the geographic cell C_k , reaching a distance

of:

$$\text{reach} - \text{dist} = \begin{cases} \text{Undefined, if } |N_\varepsilon(C_k)| < \text{MinPts} \\ \max(\text{core} - \text{dist}(C_k), \text{dist}(C_k, C_i)), \text{ Other situation} \end{cases} \quad (9)$$

(4) Redundant cells after density clustering and new geographical cells produced with time. It can calculate attribute clustering and density clustering again.

The geographical cells of big space-time data are aggregated efficiently to form a "cell-cube" structure with core geographical cells as the center. The ability of data organization and management in "cell-cube" structure is limited. It does this simply by establishing dense connections between cells. But as the cell mass increases, the limitations become more intense. The management of non-structural data and structural data in cells is chaotic, which needs more effective methods to further organize and manage them.

3.2 Spatio-temporal distribution and evolution of aggregation based on cell biomimetic structure

In the real world, the change of geographical entity belongs to the space-time change. Spatially adjacent regions with the same temporal events are conceptualized as a series of discrete and countable objects, which are called Time Feature Objects.

$$C_x = \{x(S_x, T_x) | \forall x \in S_x, T_x \in t\} \quad (10)$$

In the formula(10), S_x is the spatial position of the cell, T_x is the time when the cell is in S_x , and t is the time when the geographical cell is renewed. In the geographical space of cell biomimetic structure, the time-varying event information recorded completely represents the evolution process of geographical space. In a geospatial cube, it consists of one time dimension and two space dimensions. The time-space three-dimensional coordinate system records the change of geographical cell's time and its spatial position, as shown in figure 3. The txt documents can also be used to record changes in the specific circumstances of the geographical space, such as human causes or natural factors caused by physical changes in the geographical space. Each record, a temporal event object, corresponds to a change in a unique geographic cell.

The geographic information data is the data which exists objectively and is in the dynamic change. There are large numbers of geographical cells in geographical space. "cell-cube" also belongs to the dynamic change process under the time series. Stream data is a continuous data series under the space-time series. The macro-performance of "cell-cube" is continuous and can be affected by many factors, that is, the spatio-temporal data in "cell-cube" structure is a kind of real-time, continuous dynamic data.

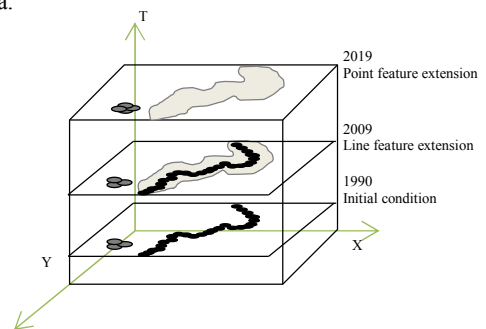


Fig. 3 Distribution and evolution of geospatial space-time

4. A SPATIAL-TEMPORAL DATA ORGANIZATION METHOD BASED ON CELL BIONICS

4.1 Construction of "cell-cube" geo-spatio-temporal data model

Traditional GIS mostly uses relational database, aiming at the cloud service mode of big data, the single relational database has some limitations in the storage and management of massive spatial data, multi-point query, association and aggregation. At the same

time, the traditional spatial database storage mostly takes the static relational data record as the main form. The management mode lacks the solution to deal with the high dynamic spatio-temporal big data. In addition to the characteristics of large volume and dynamic, the heterogeneous characteristics of large data in the existing GIS data model, there are also great difficulties in management. Different granularity of space-time, from the data format to the data storage are quite different. Common relational data structures are no longer effective in organizing and managing unstructured data. At the same time, the unified management of heterogeneous data, such as structured data and non-structured data, becomes one of the main problems that need to be solved urgently. Therefore, the "cell-cube" geo-spatio-temporal data model for big data GIS is studied to solve the problems of multi-source heterogeneous dynamics and data storage optimization. Combining the data concept and structure characteristics of geo-spatio-temporal big data and in order to solve the limitation of organization and management of "cell-cube" structure, geo-spatio-temporal big data will be divided into structured and unstructured data types. At the same time, the data feature description is extended from 3V to 5V or new 3V. By analyzing the storage and analysis problem of the leap from mass to big data, the data storage structure of stream data is described, and the data cube composed of data vectors is added dynamically according to time series. Considering the characteristics of unstructured data, the storage of geo-spatio-temporal large data needs to be scalable. Therefore, based on the horizontal two-dimensional data stream, the vertical two-dimensional unstructured data stream is added. Moreover, in the framework of time series, a "cell-volume" scalable cube model based on stream data is constructed, which lays a foundation for real-time processing of geo-spatio-temporal big data. As shown in figure 4.

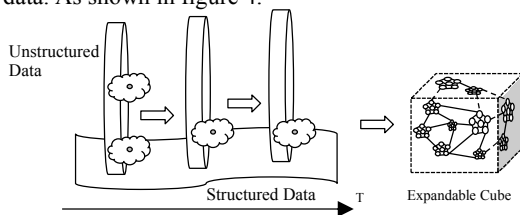


Fig. 4 The "cell-cube" extensible Cube model of geo-spatio-temporal big data

Extensible Cell-cube Model C:

$$C_{PLXYZT} = V(P,L,X,Y,Z,T) \quad (11)$$

In the formula, P is the type of information associated, L is the level of spatial resolution, X, Y, Z are the three-dimensional coordinates of the body position, and T is the time.

4.2 Storage of "cell-cube" geographical spatio-temporal data in an expandable cube

Based on the expandable "cell-cube" model, a more complete storage framework for multi-dimensional flow data is established through the process of cell biomimetic structure aggregation. The trapezoid is used to stack the flow data cubes of different dimensions. The data stream formed by dynamic data is processed by clustering according to time series. At the same time, in the vertical direction, from the bottom to the top of the cluster made regression. As shown in figure 5, The extension from the bottom layer to the top layer is the process of regression to the cluster, the higher the level, the less clustering data. The Cross section of a layer represents the cube that needs to be stored on the layer, and its volume is proportional to the number of cluster cubes stored on the corresponding regression layer.

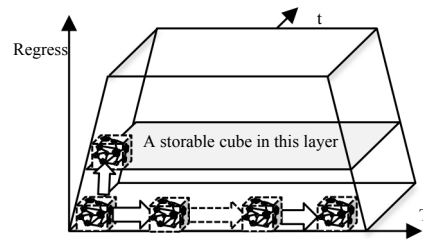


Fig. 5 "Cell-body" model storage framework based on cell biomimetic structure aggregation

Using Algorithm index-D (K, M, N,) clustering Algorithm in "cell-cube", the number of flow data cubes is the number of geographic blocks K in a certain range, each geographic block contains M geographical cells. A layer can store the geographical block for N, test factor α . In a time period of time series $T = \{t_0, t_1, t_2, \dots, t_n\}$, the clustering result is:

1) Initialize all ids to zero.

2) For each newly arrived cube data block DJ, compute clustering Streaming cube(j-1,j).

Start at Level 0 and determine if each layer has an empty storage location. Set the criteria to have at least one ID of 0 until the level where the empty cube is located.

By means of trapezoid, the data cube of horizontal flow is clustered, and the data cube of vertical flow is regressed. In the regression analysis of multi-dimensional time series flow data cube, the storage space of spatiotemporal big data can be greatly saved by using compressed regression data. In the storage structure of the cube, the one-dimensional linear regression method is used to describe the different cubes made up of different time series of the convective data cube. And the least square linear regression operation is carried out on the cubes of different time periods of the same time series. The ISB notation consists of the parameters $[Ta, Tb], \theta, f_j$. The above parameters are not correlated with each other. In the formula $[Ta, Tb]$ is the interval of time series, θ is the radix of linear fitting, f_j is the slope. Finally, in order to deal with the flow data cube in real-time, we use the method of data anomaly-driven to find the flow data anomaly cube It is feasible to make real-time response to the fast and dynamic flow data cube in the limited storage space.

Thus, considering the current dominance of relational databases, the design tools and language interfaces based on relational databases are used in the construction of Heterogeneous geo-spatio-temporal big data models Storage management for two-dimensional structured data streams in the horizontal direction. In order to solve the problem of lack of unstructured data and real-time analysis in relational database, a flexible, distributed and extended multi-dimensional flow data cube is adopted to store and manage the vertical unstructured data unstructured data makes it possible to implement both structured source and data model descriptions.

4.3 "cell-cube" geographical spatio-temporal data division

The organization and storage of data in cell biomimetic structure will be more difficult because of the increase of data amount, the processing of spatiotemporal data is slow, and the "cell-body" geographical spatiotemporal data uses the core geographical cell and its density can reach the cell attribute similarly Different from other core geographical cell attributes, data is divided to improve the efficiency of data organization and management. The degree of similarity of attributes between Core Geographical Cells C_{k_1} and C_{k_2} can be expressed as:

$$\text{Sim}(C_{k_2}, C_{k_1}) = \frac{1}{m} \sum_{j=1}^m \varphi^{a_j}(C_{k_2}, C_{k_1}) \quad (12)$$

According to the result of $\text{Sim}(C_{k_2}, C_{k_1})$, we can judge the similarity between the core geographical cells, and then judge the geographical clusters with similar attributes, as follows:

$$\varphi^{a_j}(C_{k_2}, C_{k_1}) = \begin{cases} 1 - \frac{S-1}{w}, & f(C_{k_2}, a_j) = f(C_{k_1}, a_j) \\ 0, & f(C_{k_2}, a_j) \neq f(C_{k_1}, a_j) \end{cases} \quad (13)$$

Among them,

$$S = \sum_{k=1}^w s^{a_j}(C_{k_1}, C_{k_2}) \quad (14)$$

$$s^{a_j}(C_{k_1}, C_{k_2}) = \begin{cases} 1, & f(C_{k_2}, a_j) = f(C_{k_1}, a_j) \\ 0, & f(C_{k_2}, a_j) \neq f(C_{k_1}, a_j) \end{cases} \quad (15)$$

The geographic cells of super large scale data volume are clustered into geographic clusters, and then the geographic clusters with similar attributes are divided into geographic blocks and managed in a unified way. The dynamic data can be compared with the core geographical cell attributes under the time series, the new data can be divided quickly, and the efficient organization and storage of the "cell-cube" expandable cube can be realized through the management of geographical blocks. It provides favorable conditions for compression and storage of spatio-temporal data.

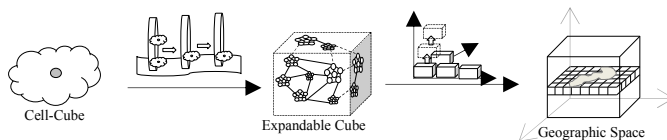


Fig. 6 "Cell-cube" spatio-temporal data model organization and management process

5. CONCLUSION

In the context of big data, the organization and management of spatio-temporal big data is a research hotspot and focus of geospatial cognition. From massive data to big data, not only has super-large-scale data volume, it also has the key characteristics of multi-source, fast, dynamic, heterogeneous and mining. In this paper, based on geographical cells and in view of the management limitations of current GIS data models, a "cell-cube" expandable cube model for geographical spatio-temporal data is proposed. Increase the vertical non-structured data cube sequence, under the time sequence, manage the structured and non-structured data of the geo-spatio-temporal big data to meet the demand of data management of geo-spatio-temporal big data with high dynamic, continuity and infinite growth. At the same time, in order to optimize the management of geo-spatio-temporal big data model, at the data level, in view of the shortage of the traditional relational database in heterogeneous and expandable aspects, the relational and non-relational databases coexist. While taking advantage of relational models for complex relational operations, massive, heterogeneous, and dynamic geotemporal big data are managed in non-relational databases, adapting to both structured and Unstructured data unified data models. Finally, on the data level, DBSCAN algorithm is used to cluster the adjacent geographical cells into geographical clusters, and then the geographical clusters are grouped into geographical blocks by the similarity of their attributes. Cell biomimetic structure can be expressed as an arbitrary three-dimensional geographical entity in efficient clustering algorithm, and the dynamic spatial cognition of spatio-temporal big data can be achieved through the updating process of geographical cells. As a kind of geo-spatial cognition based on the background of big data, the bionic structure of cells has innovated and improved in theory, and solved the problem of organizing and managing big space-time data. It is helpful to explore and mine spatio-temporal big data more efficiently.

REFERENCES

- Lu, Y.L., Li, J.W., Ye S.X., Jiang J.W., Yin M., Zhou Y.L., 2018. Big Data Organization Method of GIS Spatio-temporal Expanded the Streaming Data Cube. *Bulletin of Surveying and Mapping* (08), 115-118.
- Xu, Y.J., Tan, C.G., 2017. Study on the Spatio-temporal Data Organization and Application. *Bulletin of Surveying and Mapping* (02), 98-101.
- Li, D.R., Ma, J., Shao, Z.F., 2015. On Spatio-temporal Big Data and its Application. *Satellite Application* (09), 7-11.
- Wang, J.Y., Wu, F., Guo, J.Z., 2015a. Challenges and Opportunities of Spatio-temporal Big Data. *Science of Surveying and Mapping* 42(07), 1-7.
- Wang, J.Y., 2015b. Cartography in the Age of Spatio-temporal Big Data. *Acta Geodaetica et Cartographica Sinica* 46(10), 1226-1237.
- Hua, Y.X., 2016. The Core Problems and Key Technologies of Pan-spatial Information System. *Journal of Geomatics Science and Technology* 33(04), 331-335.

- Pei, T., Liu, Y. X., Guo, S. H., et al, 2019. Principle of Big Geodata mining. *Acta Geographica Sinica* 74(03), 586-598.
- Hua, Y. X., Zhou, C. H., 2017. Description Frame of Data Model of Multi-granularity Spatio-temporal Object for Pan-spatial Information System. *Journal of Geo-information Science* 19(09), 1142-1149.
- Qin, X., Zhen, F., 2017. Combination Between Big Data and Small Data: New Methods of Urban Studies in the Information Era. *Scientia Geographica Sinica* 37(03), 321-330.
- Jiang, N., Fang, C., Chen, M. J., 2017. Initial Exploration of Pan-spatial Cognition and Representation. *Journal of Geo-information Science* 19(09), 1150-1157.
- Wang, J. F., Ge, Y., Li, L. F., et al., 2014. Spatiotemporal Data Analysis in Geography. *Acta Geographica Sinica* 69(09), 1326-1345.
- Li, X. L., 2017. Real-time GIS Data Model Supporting Dynamic Data Management and Spatiotemporal Process Simulation. *Acta Geodaetica et Cartographica Sinica* 46(03), 402.
- Liu, H. Y., Yu, H., 2019. Multi-granularity Incremental Attribute Clustering Method. *Journal of Chinese Computer Systems* 40(03), 618-622.
- Gu, Y. H., 2018. The Methods of High Efficient Clustering on Spatiotemporal Big Data. Zhejiang, Zhejiang University, 69-71.
- Du, Q. Y., D, Z., Huang, C. D., Ren, F., 2016. Density-Based Clustering with Geographical Background Constraints Using a Semantic Expression Model. *ISPRS International Journal of Geo-Information* 5(5).
- Jing, W. P., Zhao, C. Y., Jiang, C., 2019. An Improvement Method of DBSCAN Algorithm on Cloud Computing. *Procedia Computer Science*, 147.
- K. Mumtaz, Dr. K. Duraiswamy, 2010. An Analysis on Density Based Clustering of Multi Dimensional Spatial Data. *Indian Journal of Computer Science and Engineering* 1(1).
- Li, Z. L., Liu, Q. L., Tang, J. B., 2017. Towards a Scale-driven Theory for Spatial Clustering. *Acta Geodaetica et Cartographica Sinica* 46(10), 1534-1548.
- Wu, W. H., Wu, S., 2015. Application of Density-based Clustering Algorithm in crime cases analysis considering multiple time scale. *Journal of Geo-information Science* 17(07), 837-845.
- Wu, B., 2018. Object-based Analysis and Knowledge Discovery by Modeling Spatio-temporal Evolution of Geographical. *East China Normal University*.
- Li, Q. Q., Zhang, T., Yu, Y., 2011. Using Cloud Computing to Process Intensive Floating Car Data for Urban Traffic Surveillance. *International Journal of Geographical Information Science* 25(8), 1301-1322.
- Vivoni, E. R., Camilli, R., 2003. Real-time streaming of environmental field data. *COMPUTER & GEOSCIENCES* 29(4), 457-468.
- Gong, J. Y., Li, X. L., Wu, Y. H., 2014. Spatiotemporal Data Model for Real-time GIS. *Acta Geodaetica et Cartographica Sinica* 43(3), 226-233.
- Li, Q. Q., Li, D. R., 2014. Big Data GIS. *Geomatics and Information Science of Wuhan University* 39(6), 641-645.
- Heathman, G. C., Larose, M., Ascough, J. C., 2009. Soil and Water Assessment Tool evaluation of soil and use geographic information system data sets on simulated stream flow. *JOURNAL OF SOIL AND WATER CONSERVATION* 64(1), 17-32.
- Huang, X. Z., 2015. Research on Spatiotemporal Raster Data Modeling Based on GRIC Model. *Zhejiang University*.
- Li, W. J., Yan, S. Q., Jiang, Y., Zhang, S. Z., Wang, C. L., 2019. Research on Method of Self-Adaptive Determination of DBSCAN Algorithm Parameters. *Computer Engineering and Applications* 55(05), 1-7+148.
- Liu, J. N., Fang, Y., Guo, C., et al., 2014. Research Progress in Location Big Data Analysis and Processing. *Geomatics and Information Science of Wuhan University* 39(4), 379-386.
- 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. *Bulletin of Surveying and Mapping* (04), 38-42.

ACKNOWLEDGEMENTS

The project was funded by Guangxi Key Laboratory of Spatial Information and Geomatics (15-140-07-14, 16-380-25-17).