Uncertainty in GIS spatial data

Gao, Jianxin

Jianxin Gao, "Uncertainty in GIS spatial data," Proc. SPIE 6420, Geoinformatics 2006: Geospatial Information Science, 642016 (28 October 2006); doi: 10.1117/12.712962

Event: Geoinformatics 2006: GNSS and Integrated Geospatial Applications, 2006, Wuhan, China

Uncertainty in GIS Spatial Data

Gao Jianxin*

Information Service Center, Wuhan University Library,129 Luoyu Roal,Wuhan,China,430079

ABSTRACT

A new field of technology, GIS has emerged with the development of computer technology and is now in a state of drastic evolution and development. This research discusses the present situation of the overall GIS field, the development of GIS, and features of the present development of GIS technology. It also investigates and analyzes relevant issues, such as the uncertainty in topological relations which is caused by the inaccuracy of, and uncertainty in, GIS spatial data, spatial reasoning, spatial searching, spatial-searching language, the theoretical development of topological relations, the trend of GIS development, and the problems in the development of GIS in the 21st century.

Keywords: Geographic Information System(GIS), Spatial data, Topological Relations, Uncertainty, Inconsistency

1. INTRODUCTION

Since the 1950s, the world has been swept up in a wave of technological evolution which characterizes microelectronic technology, spatial technology, information technology, and modern communication technology. And in the process of the information revolution, a brand new borderline science—geographic information system (GIS)—came to existence. GIS combines earth sciences, spatial science, environmental science, information science, and new theories of management science, as well as applied computer technology, drawing and remote sensing technology, modern geography and automatic drawing technology.

2. GIS RESEARCH COMPONETS AND RELATIONS

GIS mainly involves the description and expression of real-world spatial objects and their relations, the organization, storage, analysis, and visualization of spatial data in a computer-based environment, the design of application systems, data integration in, and business operations. It is often used to analyze and deal with the appearance and process of the distribution in a specified location, or to back up the plans or policies made by the agencies which govern natural resources and environment management. The primary mission of GIS involves analyzing spatial data, as well as accessing and transmitting spatial data. It is clear that in its forty year history, GIS has developed toward an application-and-technology-guided approach. As a result, with the deepening and widening of GIS applications, it is urgent that many fundamental theoretical issues be solved and improved. People have gradually learned that the absence of well-established geographic information theories will have serious negative impacts on the function, efficiency, and adaptability of GIS, and pose obstacles to achieving effective GIS applications and its advancement.

3. THE INDUSTRIALIZATION OF GIS AND UCGIS

To establish more theoretical support for the application and industrialization of GIS, academia, abroad or at home, has put more effort into their research on GIS basic theories, focusing on issues such as spatial relations, spatial data models, spatial cognition, spatial reasoning, and geographic information.

In the early 1980s, the research on spatial relation theories attracted much attention from GIS academia. Scholars, led by Boyle, focused on the fundamentals of spatial relations, indicating the urgency of investigating spatial relation theories and its impact on the development of GIS. Later on, the U.S. National Center for Geographic Information and Analysis listed

* jxgao@lib.whu.edu.cn; phone (+86-27) 68779089

Geoinformatics 2006: Geospatial Information Science, edited by Jianya Gong, Jingxiong Zhang,
Proc. of SPIE Vol. 6420, 642016, (2006) · 0277-786X/06/$15 · doi: 10.1117/12.712962

GIS research as one of the five priority subjects to be studied, and in its following research projects, investigations into qualitative spatial relation descriptions, perception of spatial relations in natural language, spatial and temporal reasoning, and expression of spatial knowledge. In the U.S., in 1995, a number of universities, national laboratories, and scientific technology associations set up University Consortium for Geographic Information Science (UCGIS) to carry on collaborative research in the effective access, coordination, storage, analysis, and communication of geographic information. The scope of UCGIS's research includes: ① the gathering of spatial data; ②distributed and mobile calculation; ③ the extension of geographic information representation; ④ understanding of geographic knowledge; ⑤ the interoperability of spatial data; ⑥ proportional scale; ⑦spatial analysis under the GIS environment; ⑧ future development of spatial data technology; ⑨ the uncertainty of geographic information and GIS-based analysis; ⑩ GIS and society. The research on the uncertainty of geographic information and GIS-based analysis focuses on the means and techniques involved in the spatial search, analysis, and model-establishing of GIS information. Some of the major international GIS academic conferences, such as the "International Conference on Spatial data Processing," which is held biennially, the "International Conference on Large-Scale Spatial Database," the "Spatial data Theory Conference," relevant conferences held by ISPRS (International Society for Photogrammetry and Remote Sensing), and ACSM/ASPRS conferences, have also listed spatial relations as an important part of their agenda. It is noteworthy that since the early 1990s, domestic organizations, such as the National Science Foundation, have given financial support to researches related to basic issues in spatial relations. The State Key Laboratory of Resources and Environmental Information System (LREIS), CAS. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University (LIESMARS) which delve into GIS research and application have also set up funds for studies on spatial data recognition and spatial data models; remarkable progress has been made in the research field of spatial relations.

4. THE DEVELOPMENT OF SPATIAL RELATION THEORY

Spatial relation theory covers a wide range of subjects. By and large, present research on spatial relation theory worldwide focuses on such topics as the semantic issues in spatial relations, description of spatial relations, representation of spatial relations, searching and analysis based on spatial relations, and spatial reasoning, especially the description and representation of topological relations. Researches on the abovementioned subjects have made progress and there has been a large quantity of relevant research papers and publications. However, it can be observed that no consensus has been established among the researches on the description of spatial relations, representation models and methods. All the studies were carried out under their own models. Besides, they were all performed under the premise that there is neither inaccuracy nor uncertainty in spatial data. In fact, many spatial relation researches point out in their conclusions/outlooks that the effect of inaccuracy and uncertainty in spatial data should be taken into consideration.

Meanwhile, during the past decade, the uncertainty theory of spatial information has become a special research topic of equal importance to the spatial conception mode theory in GIS. Some venerable academic institutions and universities abroad, when planning research projects in the 1990's, even set the uncertainty in GIS spatial data as one of their priority topics. For instance, topping the twelve research initiatives of NCGIA (National Center for Geographic Information & Analysis) is the accuracy of spatial databases. The regional research lab and Natural Environment Research Council of NIESR (National Institute of Economic and Social Research) proposed fourteen research initiatives, a third of which involve the "inaccurate transmission in GIS." Netherlands Spatial Data Analysis Center, a consortium of ITC(The International Institute for Aerospace Survey and Earth Science , ITC)and four universities, proposed four major research topics, the first of which, spatial analysis theory, lists the inaccurate transmission in spatial operation as its first question. Later on, more and more researches on the uncertainty in spatial data, appear in some large-scale international conferences and distinguished journals. At home, one of the fields supported by the LIESMARS established research funds is the research on the uncertainty in spatial data and the quality of spatial data. In recent years, the National Science Foundation has been subsidizing relevant researches on theories concerning the uncertainty in spatial data. For example, it subsidized the "research on the uncertainty of the position of the graphic data in GIS" in 1996; in 1998, it subsidized "the research on inaccuracy theory of the three-dimension dynamic GIS geometric information of mines", theory and methods of stochasticly expression for space entity in remote sensing and GIS integration in 2001, and "the theory of the uncertainty in spatial data and the quality control of spatial data" and "the system of spatial data mining and its application" in 2004.

It can be observed from the researches, over the past decade, on the uncertainty in spatial data, that domestic research on the uncertainty in GIS spatial data mainly focuses on the inaccuracy and uncertainty of spatial targeted location data and those of attribute data, covering issues such as: management, control, prevention of inaccuracy and uncertainty, as well as visual expressions.

GIS spatial database is mainly used to store spatial targeted location data and the data of topological relations between some targets. Most of the topological relation data derive from the known location data which is required in both application and analysis. It is obvious that using the location data which contains inaccuracy or uncertainty, as the basis for judgment, might influence the accuracy in the representation of the topological relations or result in inaccuracy and uncertainty in the topological relations, which will later impact the reliability of the spatial data searching and analysis.

Although the issue of uncertainty in topological relations has been brought up in some research literature, there is still a lack of an organized, cohesive conclusion. Generally, the research is still at a preliminary stage of exploration and this is mainly because the description, expression, and handling of the uncertainty of topological relations require theoretical support, and it is also necessary to carry out an in-depth analysis of the source, attributes, representation model, and handling methods of the uncertainty in spatial data. In other words, the theoretical support of the uncertainty in spatial data is absolutely essential.

5. THE DESCRIPTION METHOD AND MODEL OF TOPOLOGICAL RELATIONS UNDER UNCERTAINTY

In recent years, the uncertainty in topological relations caused by the inaccuracy or uncertainty of spatial data has caught the attention of the GIS field both at home and abroad, and there have been some scattered research findings, most of which focus on the topological relations between fuzzy targets. A fuzzy target is a geographic phenomenon with successive changes in distribution, such as terrain, atmospheric pressure, and commercial districts, and it is suitable for being presented with the field model. The field model can effectively present the uneven distribution of a geographic phenomenon and it uses fuzzy set theory as the mathematic tool, and attributes membership function to determine the boundary of the target. To describe the topological relations between the fuzzy targets, it first required a formalized definition of the interior, boundary, and exterior of the fuzzy target. As a result, the "egg-yolk model" was formulated. The "egg" is the maximum range of the fuzzy target while the "yolk" is the minimum range of the fuzzy target and the "white" is the scope of the uncertainty. Based on this model, forty-six RCC-model-based topological relations which divide two fuzzy regions were brought out. Interestingly, the interior, boundary, and exterior of the fuzzy target are defined in an algebraic perspective and forty-four different topological relations resulted from the use of a 9-intersection model. Besides, starting from the actual spatial data stored in the GIS database and considering a spatial target as consisting of nodes, sides, and surfaces, a representation function of the fuzzy relations between the basic elements in a 2D space is established, based on the syntactic expression. Based on Molenaar's research, Dijkmeijer et al. investigated the two preliminary topological relations between the fuzzy elements—intersection and adjacent relation. Also, Chen Xiaoyong et al. employed a 9-intersection model as their basis and used two operands in mathematical morphology: dilation and erosion, and developed a formalized dynamic 9-intersection model. Later on, a fuzzy target model was brought up, which divides the fuzzy target area into four FF-target (that is, both the interior and the boundary are fuzzy), FC-target (the boundary is fuzzy while the interior is certain), CF-target (the boundary is certain while the interior is fuzzy). A formalized description of topological relations is also presented, based on the syntactic expression of the fuzzy targets. However, Chen did not give any explicit definition of the boundary of the fuzzy target in his model.

6. APPLICATION OF TOPOLOGICAL RELATIONS TO GIS

Topological relation is the basis for GIS spatial reasoning, searching, and analysis. Many research fields of GIS, such as: spatial reasoning, spatial searching and management, and spatial searching language, are closely related to the theoretical development of topological relations. Therefore, the ways that relevant researches on topological relations develop have a direct impact on the application of topological relations to GIS. The following introduce three conditions in which topological relations are applied to GIS.

6.1 Spatial Reasoning which Integrates Topological Relations

Reasoning is a process of using known facts and rules to infer a new fact. The key problem involved in GIS spatial reasoning is how to employ the basic spatial data stored in the database, under relevant spatial restrictions, to obtain needed unknown spatial data. This involves the attributes of spatial targets and the logic expression of reasoning. The attributes of spatial targets include: topological attributes, shapes, sizes, directions, and distance, etc. Based on known data and unknown data, spatial reasoning can be divided into three categories: homogenous spatial reasoning, heterogeneous spatial reasoning, and combined spatial reasoning. A formalized method of gathering the data of topological relations is thus developed. Also, with the employment of a 9-intersection model, a combination table of the reasoning of the static topological relations is established. The table is used to answer such questions as in the following forms: "Three spatial targets—A, B, and C—are given. The topological relations between A and B, B and C are known. What is the topological relation between A and C?" An integrated representation method is developed for the reasoning of topological relations and directional relations. In addition, a combination table of the reasoning of topological relations and directional relations is developed.

6.2 The Spatial Searching Language and Spatial Searching Operation which Integrate Topological Relations

With the development of the application of computer science and technology to geographic information science, the first challenge confronted is the setup, management, analysis and operation of searching spatial databases. The spatial database contains the spatial data of the distribution of the spatial targets (data on the location and topological relations) and non-spatial data (attribute data). The searching languages of existing relevant databases, such as SQL and Quel, only provide simple operations, such as the equivalence and sorting of simple data (such as integers or characters) and fail to give effective support to the spatial searching and handling which contains spatial data. To solve problems in the application of spatial searching, analysis, and handling of spatial databases, the support of spatial searching languages is necessary. As a result, many scholars have suggested a framework of the design of spatial searching languages. However, since most researches are developed from SQL which is applied to traditional relevant databases. Setting up spatial searching, In ARC/INFO,the result that 9-intersection model describe by MARCO words (including separation, adjacent, intersection, inclusion/included, containment/contained in, equivalence) is added to the searching command, and in Oracle, the 9-intersection model is combined with SQL to expand the searching functions into space. In some literature, such SQL with spatial searching functions are called spatial SQL.

6.3 Spatial Analysis Based on Topological Relations

In a sense, spatial analysis deals with the interrelations between spatial objects. For example, overlay analysis deals with such topological relations as intersection and cover, while network analysis deals with adjacency and relation. Consistency analysis is used to examine whether a same target in multiple representations results in a topological contradiction. In the abovementioned application analyses, topological data is often considered as the most basic type of spatial data. This research investigates the consistency evaluation issues in the spatial graphic structure in multiple representations and uses the topological attributes of the spatial targets and the topological relations between the targets, as the determinants. From a spatial and temporal perspective, geographic spatial data changes with time. Such changes include not only the changes in the geometric locations, shapes, sizes of the spatial targets but also changes in the topological relations between the targets. Based on the quantitative changes in topological relation, an analysis is carried out on the similarity between the spatial graphic structures at two different times. In addition, using topological variation and geometric variation as the indexes by which to compare and analyze the data quality of two data sets which represent a same region, is one of the more important research focuses on the quality of spatial data.

(1) Uniform Representation of Spatial Data

Different representations of spatial data are built on different mathematic bases. Vector representation is the Euclidean geometry in the space of real numbers, with the attributes of point set topology. Raster representation is the discrete geometric representation in the space of real numbers, with the attributes of digital topology. Since the formalized definition and topological attributes will vary in different mathematic spaces, it is necessary to establish a spatial data representation that combines vector representation and raster representation. Based on such representation, the formalized description and division of topological relations will be carried out.

(2) Formalized Definition and Description of Spatial Targets

When a spatial target is placed in a space of another dimension, its topological attributes might change as well. However, an effective division and description of the topology of a single target is the basis for further formalized description of the topological relations between the targets. The main problem lies in current researches on this topic attempting to distinguish the topological differences between the component elements of a single spatial target and discovering how to describe the topological concepts of a single spatial target when the co-dimension is greater than zero.

(3) The Formalized Method and Model of the Description of the Topological Relations between Spatial Targets

So far, no consensus has been reached on the formalized method and model of the description of the topological relations between spatial targets, and most researches use their own method and models. Besides, most formalized models are still at the stage of roughly dividing topological relations. Existing research methods and models are still inadequate in providing detailed descriptions of topological relations, and mainly focus on the establishment, application, and calculation of topological non-variables. Those problems have greatly hindered the application of topological relations to GIS.

(4) The Description, Identification, and Management of Uncertain Topological Relations

Existing researches on the uncertainty of topological relations mainly deal with the formalized description of the topological relations between fuzzy targets, while still lacking systematic investigation of the uncertainty of topological relations, such as the cause and magnitude of the uncertainty, consistent management and attendant inaccuracy transmission models.

7. CONCLUSION

Based on the GIS spatial data, future research will focus on the topological relations under uncertainty. First of all, improvement and development should be made to the existing description models and methods of dealing with topological relations, in order to establish formalized, complete models which can be applied at varying levels. Then, a formalized description and management method of the uncertain topological relations will be developed to perfect the theoretical system of topological relations in GIS.

REFERENCES

1. G .Arbia, D .Griffith, R.Haining, Error propagation Modeling in Raster GIS.Proceedings 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences,Netherlands. 109-116 (2000).
2. Burrou Ghpa, GIS and Geostatistics:Essential Partners for Spatial Analysis. Proceedings of the International Symposium on Spatial Data Quality'99.Hong Kong : The Hong Kong Polytechnic Uni2 versity . 10-20 (1999).
3. L iDeren, On Definition,Theory and Key Technics of the Integration of GPS ,RS, and GIS. Journal of Remote Sensing. 1(1), 64-68 (1997).
4. Li Deren. Chinese Translation of the Term"Geomatics" . Acta Geodaetica et Carto Graphicasion. 27(2), 95-98 (1998).
5. L iDeren,Gong Jianya,Bian Fuling, Introduction to Geographic Information System. Beijing: Sinomaps Press, 1993.
6. Chen Jun, Zhao Ren liang, Spatial relations in GIS: a survey on its key issues and research progress .Acta Geodaetica et Carto Graphicasion. 28(2), 95-102 (1999).
7. Deng Min , Li Chengming Liu Wenbao,Describing Spatial Relations between Area Objects via Combining Topology with Metrization. Acta Geodaetica et Carto Graphicasion. 31(2), 164-169 (2002).
8. Deng Min,li Luqun,Liu Wenbao, Topological Description of Geographic 0bjects Based on Field Model within GIS. Geography and Territoria Research. 18(1), 15-18 (2002).
9. Guo Wei, Chen Jun, The Formal Description of Topological Spatial Relationship in 3D Based on Point Set Topology. Acta Geodaetica et Carto Graphicasion. 26(2), 238-244 (1997).