

# Overview on the Using Rough Set Theory on GIS Spatial Relationships Constraint

Li Jing

<sup>1</sup>College of Mobile Telecommunications, Chongqing  
University of Posts and Telecommunications  
Chongqing, China

Zhou Wenwen

<sup>2</sup>Chongqing University of Posts and  
Telecommunications  
Chongqing, China

**Abstract**—To explore the constraint range of geographic video space, is the key points and difficulties to video GIS research. Reflecting by spatial constraints in the geographic range, sports entity and its space environment between complicated constraint and relationship of video play a significant role in semantic understanding. However, how to position this precision to meet the characteristic behavior extraction demand that becomes research this kind of problem in advance. Taking Rough set theory reference involved, that make measuring space constraint accuracy possible. And in the past, many GIS rough applications are based on the equivalence partition pawlak rough set. This paper analyzes the basic math in recent years in the research of rough set theory and related nature, discusses the GIS uncertainty covering approximation space, covering rough sets, analysis of it in the geographic space constraint the adjustment range.

**Keywords**—space constraint; GIS; rough set; fuzzy geographic; spatial relationship

## I. INTRODUCTION

GIS is a spatial information system that with the capacity of the collection, storage, management, analysis and description of the earth surface and space distribution [1, 2]. And the powerful spatial analysis ability provides advantages for understanding GIS semantics, besides it becomes the crucial premise of intelligent security monitoring prototype system by using geographical boundaries. The scale is an important concept in geographical spatial cognition, also geographic process and how to extract behavior features to satisfy movement elements have a high degree of dependence to scale, therefore, to study on such problems can be converted into discussion and explore on the positioning constraints space accuracy. But in the traditional processing method of GIS data can appear error or uncertainty, which causes the result is not entirely reliable, even error, will eventually lead to the decision-making mistakes or failure. However, rough Set as a new mathematical tool of knowledge as Deal with uncertainty and fuzzy information recently has widely put into use in GIS [3, 4]. Thus it becomes a new means to measure space constraints accuracy problem.

Studies on spatial data uncertainty at home and abroad mainly focused on space position data [5] and attribute data [6], and the unsureness of spatial relationships [7], etc. The point is the study of the uncertainty modeling, precision analysis, transmission method and visual expression [8-10]. Due to the geographical phenomenon of changing over time, the spatial

data can only express one geographical phenomenon in one particular time, by that among of these studies, the considered space targets are mainly to explicitly recognize space entities (such as roads, rivers, etc.), meanwhile the uncertainty is mostly caused by measurement, digital data acquisition, and subsequent spatial analysis [11, 12]. In the approaches, to a large extent to use the methods of observational error processing in geodesy [13]. The vagueness refers to the thing can be clear described by the predefined attributes, but the boundary of the adjacent target is hard to differentiate clearly. Thus, the rough set theory has been widely used in the field of image processing, which includes system simplification, remote sensing image segmentation and remote sensing image recognition, etc. [14, 15].

This paper mainly probes into the uncertainty of this kind of spatial constraint, which includes introducing rough set into the geographical border uncertainty research, besides its basic concepts and essence (in part 2), the expression methods and rules of spatial relationship, and the applied analysis in GIS (in part 3), finally this paper discusses the current research progress, future development and difficulties (in part 4 and part 5).

## II. ROUGH EXPRESSION AND REASONING OF SPATIAL RELATIONSHIP

### A. Basic conception of rough set

Rough set theory is a mathematical tool proposed by professor Pawlak in 1982 which can quantitative analyze and deal with imprecise, inconsistent, incomplete information and knowledge [16]. Rough set theory, the initial prototype from the relatively simple information model, its basic idea is formed by classified concepts and rules of the relational database, through the classification of the equivalence relation and classification for the target approximate knowledge discovery. And the following are four kinds of rough set basic definitions methods [17, 18]:

#### 1) Equivalence relation and the indiscernibility relationship

Set  $R$  is the equivalence relation limited on the  $U$  field (Meet the reflexivity, symmetry, and transitivity), recorded as  $R \subseteq V \times U$ . In  $U$  field, all  $x \in v$  has the equivalent relation with the collection of  $R$ , recorded as  $[x]_R = \{y \in U \mid (x, y) \in R\}$ .  $U/R$  indicates all the sets of  $R$  constitute of the equivalence classes that is quotient set.

And set R is as the equivalence relation clan, in which  $P \subseteq R$ , and  $P \neq \emptyset$ . All the intersection of equivalence relations in P is called the indistinguishable relationship of P, which can be recorded as  $\text{ind}(P)$ , and that is  $[x]_{\text{ind}(P)} = \bigcap (R \in P)[X]R$ .

2) *The upper approximation, lower approximation and rough sets*

The uncertainty of rough set theory is based on the concept of upper and lower approximation. Set R is the equivalence relation limited on the U field, considered to the set  $X \in U$ , coupling  $(RX, \overline{RX})$  is called a rough approximation of X in the approximate space  $(U, R)$ , and

$$RX = \{x \in U \mid [x]R \subseteq X\}, \overline{RX} = \{x \in U \mid [x]R \cap X \neq \emptyset\} \quad (1)$$

In which  $RX, \overline{RX}$  indicate the R upper and lower approximation of X respectively. Also the set  $\text{bn}R(X) = RX - \overline{RX}$  refers to the R boundary region of X. The rough set defines that when  $\text{bn}R(X) = \emptyset$ , means  $RX = \overline{RX}$ , it refers to that X is R accurate sets and when  $\text{bn}R(X) \neq \emptyset$ , means  $RX \neq \overline{RX}$ , it calls X is R rough sets.

3) *Variable precision rough set definition*

The variable precision rough set definition is the extension of Pawlak rough set, it introduces  $\beta$  ( $0 \leq \beta < 0.5$ ) to the basic rough set, therefore it should define the majority incorporate coefficient  $\beta$ , before the variable precision rough set.

Definition: Set X and Y are the non-void subset of the limited discourse domain U. If to each  $e$  ( $e \in X$ ), there is  $e \in Y$ , and called it Y includes X that is recorded as  $X \subseteq Y$ . And also make

$$C(x, y) = \begin{cases} 1 - \frac{|X \cap Y|}{|X|} & |X| > 0 \\ 0 & |X| = 0 \end{cases} \quad (2)$$

In which  $|x|$  is the cardinal number of set X and  $C(x, y)$  is the relative error resolution of set X regarded to set Y.

Definition: Based on the majority incorporate coefficient relationship, set  $(U, R)$  as approximation spaces, in which discourse domain U is non-empty limited set, and R is the equivalence relation on U,  $U/R = \{E_1, E_2, \dots, E_n\}$  is the set constituted of equivalence class or basic set of R. According to  $X \subseteq U$  define the  $\beta$  lower approximation  $\underline{R}_\beta X$  and the  $\beta$  upper approximation  $\overline{R}_\beta X$  of X, and it approaches to variable precision rough set definition as:

$$\underline{R}_\beta X = U \{E \in U/R \mid c(E, X) \leq \beta\} \quad (3)$$

$$\overline{R}_\beta X = U \{E \in U/R \mid c(E, X) < 1 - \beta\} \quad (4)$$

4) *Reduction*

Set U as a discourse domain, Q and P are defined as two equivalence relation clusters on U besides  $Q \subseteq P$ , if Q is independent and  $\text{ind}(P) = \text{ind}(Q)$ , then call Q is the absolutely reduction of the equivalence relation cluster P and recorded it as  $\text{red}(P)$ . All the absolutely relationship sets in P are the core of equivalence relation cluster P, recorded as  $\text{core}(P)$ .

The most significant difference of rough set theory and other theories of dealing with uncertainty and imprecise problems is about processing not provide any prior information

except data collection, so the description of the uncertainty or processing can be said to be more objective [19], therefore it provides preferential conditions for the study of spatial direction relationship.

B. *The Rough thoughts of spatial direction relationship*

If it is necessary to integrate the rough set theory to the GIS, must start from the basic data model to establish uncertain, fuzzy geographic data model, so as to solve the two kinds of inaccuracy and fuzzy problems between the direction relationships of the objects, and then solve the uncertainty caused by fuzzy object boundary problems. In the spatial relationships [20], there are fuzzy and precise objects in the space object, therefore, the spatial direction relation can be mainly divided into four types: fuzzy objects and the direction relationship between fuzzy objects, the fuzzy objects and the direction relationship between precise objects, the direction relationship between the fuzzy and precise objects, and also precise objects and the direction relationship between the precise objects. Because the space objects can approximately be expressed by the rough set, the spatial direction relationship between fuzzy and precise objects can be solved by researching on the relationships between its upper and lower rough approximations sets.

Set the fuzzy objects A and B, the upper and lower rough approximations sets of them are  $RA, \overline{RA}, RB, \overline{RB}$  separately, and adopt 048 to represent the direction relationships of them. When the grading of the direction relationships is as 8, the direction relationship knowledge base is regarded as  $\{N, NE, E, SE, S, W, SW, NW, O\}$ , when the grading is four direction relationships, the knowledge base is regarded as  $\{N, E, S, W, O\}$ . The rough expression of the spatial objects direction relationships is to represent the concept of complex direction relations as a collection of basic knowledge of the knowledge base. In which the classification knowledge of the discourse domain space is known, and the key is to confirm the relational functions between the concepts and the basic knowledge. There are two kinds of accuracy and fuzziness issues about direction relationships of the objects. The first one is caused by the fuzzy boundaries of the fuzzy objects, and the second one is by the improper methods which are caused by adopting the basic direction relationships of the knowledge base to represent the object direction relationships. The former is inherent, and the latter is the issue of methods and cognition.

The rough expression of spatial direction relationships. Set the extension cord of the outside rectangular of the object A divide the space region into  $O_i$  ( $1 \leq i \leq n$ ), and n is as the resolution ratio of the direction. The membership function of objects B and  $O_i$  is as:

$$U(B \in O_i) = \begin{cases} 1, & B \cap O_i \neq \emptyset \\ 0, & B \cap O_i = \emptyset \end{cases} \quad (5)$$

The direction relation of objects A and B is  $O_{AB} = \{O_i \mid U(B \in O_i) = 1\}$ . Another definition is: the upper and lower rough approximations sets of the fuzzy objects direction relationships  $O_{AB}$  are  $\underline{O}_{AB}$  and  $\overline{O}_{AB}$ ,  $\underline{O}_{AB} = \{O_{AB}\}$ ,  $\overline{O}_{AB} = \{O_{AB}, O_{AB}, O_{AB}\}$ .

Because the fuzzy and precise objects can be unified expressed by the upper and lower rough approximations sets,

therefore by adopting the rough set is with the ability to unify the approximately express of the direction relationships of the fuzzy and precise objects in the frame, in order to process and reason. The upper and lower rough approximations sets of the fuzzy and precise objects direction relationships are equal, so the boundary is empty, it is consistent with the traditional expression method which eight direction relationship is based on projection, the approximation precision is 1. The boundary of the rough expression of spatial direction relationships is mainly caused by the boundary of the fuzzy objects, so the method can describe the fuzziness direction relationships of the which are caused by the fuzziness of the fuzzy objects boundary, it mainly is to approximately express the fuzzy direction relationship created by the fuzzy objects boundary, but it fail in solving the second kind of inaccuracy and fuzzy issues. [21]

### III. GEOGRAPHY SPATIAL RELATION RULES EXTRACTION BASED ON ROUGH SET

#### A. Rough set expression of spatial relationship

To quantitatively express all kinds of the geological phenomenon spatial relationship, and then effectively converted into the format of the rough set data processing method, is the necessary conditions to use rough set rules to extract major spatial relationship of geological phenomena. Because of the reason that the rough set need to represent the data into the form of a two-dimensional table as processing the data, accordingly it requires to various kinds of geological phenomenon spatial relationship of two-dimensional form.

##### 1) choose spatial relationship:

To aim at the specific issues of the geography, it chooses the specific spatial relationships of the geological phenomenon as the research objects according to the prior knowledge. Such as the respective features of different geological phenomenon: the water cycle, atmospheric circulation, ocean vortex, land usage and coverage, select the major effected spatial relationship factors such as the distance, topology and etc.

##### 2) quantitative expression of spatial relationship:

To aim at the various spatial geological phenomenon, it adopts the appropriate description methods to quantitatively describe the spatial relationships, for example by employing the Euclidean distance to quantitatively describe the distance.

##### 3) Construct spatial relationship decision table:

To convert quantitative description of geological phenomena spatial relationships into the form of a decision table. And the rows of the decision table say the research objects of geological phenomena, on the other side the columns of the decision table represent two parts: The former part known as condition attributes, on behalf of all kinds of geological phenomenon spatial relationships, the latter part of the decision table is decision attributes, the values of them are specific geological results. The values of each row are the quantitative descriptions of spatial relationships approached by various description methods of spatial relationship (except decision attributes). By using this two-dimensional table to express the spatial relationship of geological phenomena, we can employ the method of the rough set to analyze and extract

the main spatial relationship rules of the geological phenomenon.

#### B. The spatial relationship rules extraction

Using the rough set method to process the geological phenomenon of intrinsic spatial relation rules extraction, it is mainly divided into the following steps:

1) The spatial relationship of rough sets expression: aim to the study of geological problems, by the method in III (A), do the processing of expressing the geological phenomena of spatial relationship to the data processing format of the rough set -- the form of a decision table.

2) Using the discretization method of the rough set theory to get the decision table then to discretize. As the rough set to process the decision table, it requests the values in the decision table expressed by discrete data (such as integer, string type, enumeration type), therefore, before processing the data it must do the decision table discretization.

3) Using the attribute reduction algorithm of the rough set to do the processing of spatial relationship reduction on the space relationship the discrete decision table of the geological phenomenon space relationship, and finally form the space relationship decision rule table. The spatial relationship decision table after reducing then become the space relationship decision rule table. Because the results of the space relationship reduction are not unique, and each reduction result of the space relationship decision table will become one space relationship decision rule table, so the finally, space relationship decision rule table is the "and" of all the space relationship decision rule tables that came from each reduction result. To the final spatial relationship rule, that asks for calculating the coverage and confidence of the spatial relationship decision rules.

### IV. ROUGH SET THEORY IN THE APPLICATION OF GIS SPATIAL RELATIONSHIPS

#### A. GIS data

Data analysis is an important part of GIS data processing. Rough set theory has some unique opinions such as knowledge granularity, new membership, which makes rough set particularly suitable for data analysis, therefore, there are some successful applications by using the rough set theory in GIS data analysis, for example, to adopt Worboys to handle the inaccuracy caused by multi-space or multi-semantic resolution ratio [22, 23] models like Theresa based on rough set and Egg-Yolk model study on the fuzzy and uncertainty problems of spatial data [24]; Du introduces the rough set theory into the expression of direction relationships, and present the direction relationship rough expression method, variable precision rough representation methods and rough reasoning method of direction relationship, which leads to enhance the processing and analysis ability to handle accuracy and fuzziness, and also can unify the direction relationship between the fuzzy objects and the precise objects into a framework [25]; Shi has already discussed on the rough set theory in the application of GIS uncertainty problems, which shows the rough set theory is

valuable in GIS uncertainty, but recent researches have not get deeply [26].

### B. Spatial data mining

GIS is the main part of the spatial database development and contains a large number of spatial and attribute data, which has more rich and complex semantic information than the general database, and hides abundant information, all of these are very necessary for data mining. Spatial data mining means to extract the information users interested in which includes common relationships of spatial patterns and features, or spatial and non-spatial data, and some other general data characteristics hidden in the database data. Accordingly spatial data shows increasing important in the found and remake nature projects of people activity, the research and application of spatial data mining also increasingly aroused concerns, and the rough set theory is one of the important methods introduced to the data mining, in 1995 Theresa Beaubouef tried to describe a database model based on the original rough sets theory, and introduced some rough relational database models which include systems involving ambiguous, imprecise, or uncertain data [27], and Wang used GIS attribute mining as an example to analyze the application of rough set in GIS data mining [28].

### C. Fuzzy geographical object modeling

The fuzzy object modeling have a wide range of meanings. The real world is complicated and full of all kinds of uncertainty, however in GIS, traditional geographic object modeling only consider the clear objects cannot reflect the complexity and uncertainty of the real world well. That causes the poor decision ability of GIS based on these kind models, which leads to hinder the development of GIS intellectualization. Using rough set to describe fuzzy object, is with the ability to fully represent the fuzziness of fuzzy objects, therefore abundant researches and applications of geographic object modeling based on the rough set theory have emerged, such as the research of Liao [29] is based on the rough set theory to transfer method to consider the polygon boundary of data fuzziness, and to employ the membership function to determine the uncertainty of the polygon boundary. Besides Du [30] combined the advantages of the rough sets and fuzzy sets dealing the fuzziness and uncertainty of the spatial data to express the fuzzy objects, and leads to expand the space data model expressing ability of fuzzy data.

### D. The combination of rough set and other soft computing methods

The rough set theory is one kind of soft computing method, and the purpose of the soft computing method is to adapt to the inaccuracy of the real world around, to explore the tolerance to the accuracy, the uncertainty and partial real, and in order to achieve hand lability, robustness, and better contact with reality, whose function model is the mind of human. The main methods of calculating software are with rough sets and fuzzy sets, neural networks, genetic, and the theory of transport, etc. As solving practical problem, to adopt several computing technologies collaboratively rather than mutually exclusively has superiority compared with using one kind of computing technology. And also, it can combine the various sources of knowledge, technology and methods which solving complicated practical problems ask for. Due to the rough set

has certain shortcomings as processing the data, it is necessary to combine the rough set method with other uncertain methods. At present, there are some applications of GIS data processing that combined the rough set with other soft computing methods, and the more commonly used is the combination of rough set and neural network or fuzzy sets [31, 32].

## V. CONCLUSION

Rough set theory is a data analysis tool, which provides a powerful tool for the expression of GIS uncertainty information and processing, and offers favorable conditions to solve uncertain boundary space constraints. In which the fuzzy set and probability statistics method are also the commonly used methods dealing with uncertain information, but these methods need some additional information or data prior knowledge, such as fuzzy membership function and probability distribution, however sometimes it is not easy to get the information. On the other hand the rough set theory just use the information provided by the data itself, without any prior knowledge, at the same time has great advantage to reveal and express multi-level spatial knowledge.

To make a better use Rough set theory in GIS, there are still many problems to be solved. Mainly displays in: rough set can only be used for discrete space, and must be qualitative, therefore only apply to raster data, the application of vector data is difficult to determine; Rough set theory to study the expression of uncertainty in spatial analysis: recently rough set is used in attribute data, involved little in the location data uncertainty. Combining rough set and other uncertain methods, it although has made some achievements, but still there is a lot of unsolved problems ask for further research. With the further increasing of GIS data processing requirements, rough set theory is widely used to spatial data processing, at the same time, it will promote the development of the future GIS data processing technology, especially the spatial decision support system.

## REFERENCES

- [1] Zhang Xiao-Xiang, Yao Jing, Li Man-Chun, A Review of Fuzzy Sets on Spatial Data Handling[J], Remote Sensing Information, 2005(2).
- [2] Robinson V B. Some implications of fuzzy set theory applied to geographic databases [J]. Computers Environment and Urban Systems, 1988, (12):89-97. doi:10.1016/0198-9715 (88) 90012-9.
- [3] Liu Wenbao, DengMin, Analyzing Spatial Uncertainty of Geographical Region in GIS [J], Journal Of Remote Sensing, 2002, 6(1).
- [4] LIAO Wei-hua, Method Study of GIS Data Transformation Based on Fuzzy Rough Set[J], Remote Sensing Technology And Application, 2007, 22(6)
- [5] (Zhang Jingxiong Du Daosheng, Field-based Models for Positional and Attribute Uncertainty [J], Acta Geodaetica Et Cartographica Sinica, 1999, 28(3).
- [6] Shi Wenzhong, Wang Shuliang, State of the Art of Research on the Attribute Uncertainty in GIS Data[J], Journal Of Image And Graphics, 2001, 6(9)
- [7] CHENG Jicheng, JIN Jiangjun, The Uncertainty of Geographic Data[J], Geo-Information Science, 2007, 9(4)
- [8] WANG Xiaoming, LIUYu, ZHANGJing, Geo -Spatial Cognition : An Overview [J], Geography and geographic information science, 2005, 21(6), pp. 1-10.
- [9] Roy AJ, Stell JG. Spatial relations between indeterminate regions. International Journal of Approximate Reasoning, 2001, 27(3), pp. 205-234.

- [10] Leung Y, Ma J H , Goo dchild M F. A General Framework for Error Analysis in Measurement based GIS [C]. The 2nd International Symposium on Spatial Data Quality, Hong Kong, 2003.
- [11] Hu Shengwu, Wang Hongtao, Representation and Properties Researches about Fuzzy Geographic Entities[J], *Geomatics & Spatial Information Technology*, 2006,29(2).
- [12] Jonathan Lee, Introduction: Extending Fuzzy Theory to Object-Oriented Modeling[J], *International journal of intelligent systems*, 2001, 16(7).
- [13] Cheng Tao, Deng Min, LI Zhilin, Representation Methods of Spatial Objects with Uncertainty and Their Application in GIS[J], *Geomatics and Information Science of Wuhan University*, 2007, 32(5), pp.389-393.
- [14] Sun Lixin, Gao Wen, Selecting The Optimal Classification Bands Based On Rough Sets[J], *Pattern recognition and artificial intelligence*, 2000, 13(2).
- [15] XU Yi, LI Longshu, Image Segmentation Based on Rough Entropy and K-Means Clustering Algorithm[J], *Journal Of East China University Of Science And Technology(Natural Science Edition)*, 2007, 33(2).
- [16] Pawlak Z. Rough set. *International Journal of Computer and Information Sciences*, 1982(11), pp.341-356.
- [17] Han Zhenxiang, Zhang Qi, Wen Fushuan, A Survey on Rough Set Theory and Its Application[J], *CONTROL THEORY & APPLICATIONS*,1999, 16(2).
- [18] Pawlak Z. *Rough set-theoretical aspects of reasoning about data*[M],.Dordrecht:Kluwer Academic Publishers,1991
- [19] Wang Guoyin, Yao Yiyu, YuHong, A survey on rough set theory and applications[J], *Chinese journal of computer*, 2009, 32(7), pp. 1229-1246.
- [20] Liao Weihua. GIS uncertainty analysis based on covering rough set [J], *Science of Surveying and Mapping*. 2012, 37(4), pp.154-156.
- [21] Burrough P A, Frank A U. *Geographic Objects with Indeterminate Boundaries*[M].Basingstoke:Taylor and Francis,1996.
- [22] WORBOYS M. computation with imprecise Geospatial Data[J].*Computers, Environment and Urban Systems*,1998.85-106.doi:10.1016/S0198-9715(98)00023-4.
- [23] WORBOYS M. Imprecision in Finite Resolution Spatial Data [J].*Geo Informatica*,1998,(03):257-279.).
- [24] THERESA B, FEDERICK E P. Vague regions and spatial relationships: A rough set approach [A].2001.pp.313-318.
- [25] Du Shihong, WangQiao, Spatial Orientational Relations Rough Reasoning[J], *Acta Geodaetica Et Cartographica Sinica*, 2003, 32(4) : 334—338.
- [26] Wang Shuliang, Li Deren, Theory and Application of Geo-rough Space[J], *Geomatics And Information Science Of Wuhan University*, 2002, 27(3), pp. 274—282.
- [27] Theresa Beaubouef, Frederick E. Petry, Bill P. Buckles, Extension Of The Relational Database And Its Algebra With Rough Set Techniques[J], *Computational Intelligence*, 1995, 11(2),pp.233-245.
- [28] Deng Xueqing, Dong Guangjun, Gis Attribute Data Mining Based On Rough Set Theory[J], *SURVEYING AND MAPPING OF SICHUAN*, 2003, 26(4) .
- [29] LIAO Wei-hua, Method Study of GIS Data Transformation Based on Fuzzy Rough Set[J], *Remote Sensing Technology And Application*, 2007, 22(6)
- [30] Du Shihong, WangQiao, The Reserch of Rough Expression of Fuzzy Objects and their Spatial Relations[J], *Journal Of Remote Sensing*, 2004, 8(1)
- [31] Greco, S., Matarazzo, B.S., S lowinski, R., "Rough membership and Bayesian confirmation measures for parameterized rough sets", *Rough sets, Fuzzy Sets, Data Mining, and Granular Computing, Proceedings of RSFDGrC'05, LANI 364 1*, PP. 314 - 324, 2005
- [32] Ali Azadeha, Morteza Saberi, An integrated Data Envelopment Analysis–Artificial Neural Network–Rough Set Algorithm for assessment of personnel efficiency[J], *Expert Systems with Applications*, 2011, 38(3), pp. 1364–1373.