

## CARDINAL DIRECTION RELATIONS QUERY MODELING BASED ON GEO-ONTOLOGY

Xinyan Zhu<sup>a</sup>, Di Chen<sup>a\*</sup>, Chunhui Zhou<sup>b</sup>, Ming Li<sup>a</sup>, Weidong Xiao<sup>c</sup>

<sup>a</sup> State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, P.R.China - geozxy@263.net, geoehen1202@gmail.com, Liming10307@163.com

<sup>b</sup> Navigation College, Wuhan University of Technology, 122 Luoshi Road, Wuhan 430063, P.R.China - church.zhou@gmail.com

<sup>c</sup> C4ISR Technology National Defense Science and Technology Key Lab, National University of Defense Technology, Changsha 410073, P.R.China

### Commission II, ICWG II/IV

**KEY WORDS:** Direction relations, Query Model, Digital gazetteer, Spatial reasoning

### ABSTRACT:

Direction relations, as an important spatial relationship, is simply expressed as object prosperity in traditional geo-ontology. The lacking of explicit specifications and reasoning rules of direction relations in geo-ontology result in the difficult or inflexible of spatial reasoning. Also, digital gazetteers provide information on named features, linking the feature's name with its location and its type. Although the location information is incomplete and not exact, the implicit spatial information, for example spatial relationships and spatial scale, can be extract using the appropriate models based on geo-ontology. In this paper, we proposed a novel conceptual framework of direction relations in order to formalize the semantics and implicit information of direction relations, and present an extraction algorithm of implicit information based on previous researches, which will produce a complete query instance of direction relations. At last, the most suitable direction physical model is recommended to calculation module according to relevant rules. And the experimental results show that this direction query model not only extracted the implicit information effectively, but also made a reasonable interpretation for the user's intention.

### 1. INTRODUCTION

The traditional conceptual model of directions can be denoted by *<reference region, target region, direction>*. Take two following queries for example, Q1={“Provinces lie to the north of Changjiang river”} and Q2={“Cities lie in the north of Hubei province ”}. These queries will be mapped to traditional conceptual model of direction relations by conceptual matching, which can be split into two triplets, T1: *<Changjiang river, Provinces, north>* and T2: *<Hubei province, Cities, north>*. It seems that the different between them is small except geographic objects. But there are many factors to affect direction relations such as geometry type, relative scale, shape, etc. From perspective of spatial scale, Q1 described the direction relations between polygon and polyline in large-scale space, while Q2 described the direction relations between polygons in small-scale. It is obvious that the direction description models (physical models) established for these two queries are different, according to the geometry types and the spatial scales of reference objects and target objects. For traditional models many influence factors, which represent people's cognitive custom, are taken into account directly in physical level. In fact, the difference of spatial cognition have exist and even been magnified in the process of mapping from semantic tier to physical tier, and often result in the calculated results are not good to meet the query requests submitted by users. It is necessary to deeply analyze the influencing factors and to extract some implicit information as much as possible.

Geo-ontology, which address the issue of lacking semantic, offer a formal, explicit specification of a shared

conceptualization to geographic field (Wang, 2007). The geographic semantics and the relationships between concepts can be obtained, but they are ineffective and inflexible to support complicated spatial reasoning, especially to spatial direction relationships. And existing direction conceptual model or query model based on geospatial semantics omit some implicit information, including the relative scale between objects and the fuzzy degrees of direction concepts, and do not choose the most appropriate physical model automatically depend on the actual situation to perform the detailed calculation. Although various physical models have been proposed such as projection-based, cone-based, triangular model, direction relations between MBRs, direction-relation matrix, voronoi diagram-based, statistical model, etc (Goyal, 2000; Du, 2008; Deng, 2008), these models are appropriate for different environments because of the vagueness and complexity of direction relations, for instance, considering the different type and spatial scale of objects and the degree of vagueness of direction description. Considering the limitations of existing query model, we proposed a novel conceptual model of direction relations in order to formalize the implicit semantics of direction relations in order to addressing the semantic differences between user's original intention and physical model. And we integrate various direction description models as candidate models which will be chosen to calculate direction according to the characteristics of geographical object. The aim of this query model is to convey user's intention from semantic tier to physical tier as full as possible via extracting the implicit information. Then a valid result can be obtained through the reasoning and calculation on physical tier.

\* Corresponding author. Email Address: geoehen1202@gmail.com

The remainder of this paper is organized as follows. In Section 2, we proposed a conceptual model of spatial direction relations based on query characteristics. In Section 3, an algorithm flow of extracting implicit information is presented. Then, some experiments prove that our query model is feasible and accurate based on geo-ontology and gazetteer in Section 4. Finally, Section 5 concludes the paper and points out the possible future research.

## 2. CONCEPTUAL MODELING FOR DIRECTION RELATIONS

There are many different representations of the same geographic reality for the difference in spatial cognition or to meet the particular application needs. The pertinent tests show that human cognition of spatial direction is closely related to their regions, backgrounds and cognitive environments (Jin, 2010). Thus, we are trying to acquire the user's understanding of geographical direction or query restrictions on the semantic tier.

### 2.1 Direction Query Characteristics

As the layer of commitments mediates between the ontology base and applications, there has been increasing recognition that contexts are important issues in semantic interoperability. In the process of query modelling, the original intention is gradually simplified and formalized, and a lot of useful information for building the physical model of direction relations is omitted or hidden with a simple query expression, including the distance between objects, relative scale and geometry type of objects and so on. When these factors, referred as *direction query characteristics*, are relatively definite, the physical model of direction relations can also be determined. But it is noteworthy that the more direction query characteristics define, the larger deviation between the calculation model and the real meaning of the questioner. It is not guaranteed that each query characteristics considered conformed to user's intention because the potential errors have always been for each of them. This is exactly what the embodiment of the difference and vagueness of spatial cognition. Therefore, some of the major influencing factors are chosen as the query characteristics and the result of information extraction, including geographical range of research, geometry type, geography type and the properties of geographical object and reference system of cardinal direction.

### 2.2 Conceptual Model of Spatial Direction Relations

In fact, two main steps are required during the mapping from user queries to concrete physical model. The first is to enrich and restore the simple query expression from simple to detail, which can be regard as an inverse process of query modelling; the second is to construct a specific physical model and perform geometric calculations according to the extended query expression. However, the key of mapping process is how to specify and formulate the extended query. So we proposed a novel conceptual model of direction relations for establishing the link between context tier and semantics tier. It can be expressed as follows:

#### Direction Relations

$::= \langle \text{ref\_Region}, \text{tar\_Region}, \text{ref\_Frame}, \text{qDir} \rangle$

#### ref\_Region | tar\_Region

$::= \langle \text{geographicClass}, \text{geometryType}, \text{rel\_Scale} \rangle$

$\text{geometryType} ::= \text{point} \mid \text{line} \mid \text{polygon}$

$\text{ref\_Frame} ::= \langle \text{dir\_Num}, \text{topo\_Ref}, \text{parti\_Form} \rangle$

$\text{topo\_Ref} ::= \text{Interior} \mid \text{External} \mid \text{Boundary} \mid \text{Intersect}$   
 $\text{parti\_Form} ::= \text{Project} \mid \text{MBR} \mid \text{Cone-shape} \mid \text{Others}$   
 $\text{qDir} ::= \langle \text{dir}, \text{cons\_Degree} \rangle$   
 $\text{dir} ::= \text{N} \mid \text{NE} \mid \text{E} \mid \text{SE} \mid \text{S} \mid \text{SW} \mid \text{W} \mid \text{NW} \mid \text{Mid}$

*ref\_Region* and *tar\_Region* represent reference region and target region, they all have three slots: *geographicClass*, *geometryType* and *rel\_Scale*. The first slot indicates the geography class of spatial objects in geo-ontology; the second indicates the geometric type of object, such as point, polyline and polygon; the third indicates the relative spatial scale between reference region and target region, and go into detail in next subsection.

*ref\_Frame* denotes the physical direction model that will be used to calculate and describe direction relations, and also have three slot: *dir\_Num*, *topo\_Ref* and *parti\_Form*. The first indicates the number of cardinal direction, and its value range is 2, 4 or 9. They are  $d_2 = \{W, E\}$  or  $\{N, S\}$ ,  $d_4 = \{N, E, S, W\}$  and  $d_9 = \{N, NE, E, SE, S, SW, W, NW, O\}$ . The second indicates the topological relationship between reference region and target region, which are three major types (*Interior*, *External* and *Boundary*); the third indicates the parting form of research region, such as projection-based, cone-based or MBR-based or others, it is associated with geometry type of objects and *dir\_Num*.

*qDir* denotes the specific direction that described in user's inquiry. The *dir* slot is the direction name; the *cons\_Degree* slot defines the degree of constraint of direction concepts, i.e. the extent of direction constraint of *due north*, *north by east* and *northeast* are different. So each direction concept is attached a value by adding a property of *hasConstrainDegree*, and its values range from 0 to 1. It has a direct impact on the process of establishing physical model and the fuzzy calculation and description.

And Figure 1 shows the conceptual model of direction relations, Protégé is selected to create OWL DL expression of ontology knowledge base. Take Q1 for instance, this query can be mapped to this expanded conceptual model using the Conceptual Matching Algorithm (Mata, F. 2007), yet it is a more coarse query instance because some slot values require geometry calculation and spatial reasoning. The result of processing is as follows:

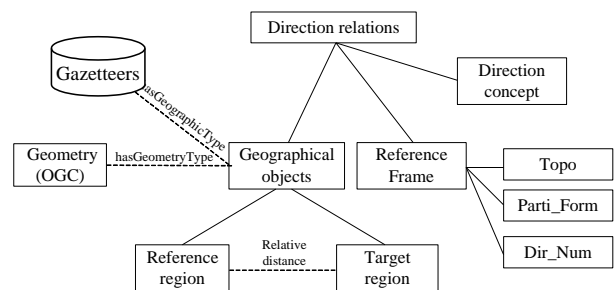


Figure 1. The implementation of conceptual model

$\langle \text{River}, ?, ? \rangle, \langle \text{Province}, \text{Polygon}, ? \rangle, \langle ?, ?, ? \rangle, \langle \text{north}, 0.75 \rangle \rangle$

Where “?” express the slots that is unknown or remain to be deduced with inference rules. Therefore, the goal of next section

is to determine these slots value using a series rules base on existing gazetteers and geographical ontology.

### 3. EXTRACTION RULES FOR QUERY CONTEXT BASED ON GEO-ONTOLOGY

#### 3.1 Footprint in gazetteers

Digital gazetteers are directories containing triples of place names (N), geographic footprints (F), and feature types (T) for named geographic places (L. Hill, 2000). And they have play a key role in geographical information retrieval, But existing gazetteers are not employed since they lack the capabilities to support logical inference (Carsten K, 2010). To achieve the complex query of direction relations base on digital gazetteers, it is important to make full used of ontology semantic information for reasoning like feature types and categories of places and geographic footprint representing the location of a named place for detail spatial calculation. And footprint might be point, bounding box, line, polygon and grid representation. It is possible that the identical place can have multiple footprint representations. Place names are inherently imprecise and each footprint used to describe a place is to some degree an approximation (L. Hill, 2000). Each geographical feature possesses an object property *hasGeometryType* to state its specific geometry type in gazetteers, which are organized in accordance with the ontology structure. For example, River *hasGeometryType* polyline. But footprint is a simplified shape of geographical feature. The footprint may be a single point for small linear or polygon feature. In fact, footprint can be seemed as the abstract and generalization of place. On the other hand, most existing models describe the direction relations of approximations rather than the real geospatial objects. The deviation and uncertainty will directly influence the correctness and precision of spatial reasoning result. Therefore, in the next step, which is to improve and complete the generated coarse query instance mentioned before, it is essential to consider the uncertainty and fuzziness in detailed calculation based on footprint.

By Tobler's First Law of Geography known, near things are more related than distant things. In other words, the farther apart are two geographical objects, the more blurring are their shape from the spatial cognition perspective. The accurate shapes of geospatial objects are not important in the process of calculating and describing the direction relations, because the effect decreases gradually with increasing distance. Thus they could be abstracted or generalized into simply or low dimensional geometrical objects. Meanwhile, the direction relations between two geographical objects can keep consistent in different spatial scale (Du, 2010). It is reasonable and necessary to used approximations to perform query processing considering the complexity of calculation and fuzziness of cognition. In previous research (Chen, 2011), we proposed a fuzzy description model of direction relations and some relevant concepts, such as Relative Distance and Sampling Granularity model. And the greater relative distance between reference region and target region, the larger sampling granularity (abstract degree). For example, a polygon feature can be abstracted as a point. And the degree of generalization can be quantitative using relative distance which is depended on the distance between two objects and their size and shape. Although footprint in digital gazetteer is imprecise and rough, it is practical to calculation direction relation from footprints. For the place that has multiple footprint representations, which kind

of geometry expressions should be chosen for geometry calculation is not only related to geographical type, but also to relative distance between objects.

#### 3.2 Extraction algorithm of query characteristics

Though the abstraction or generalization is to simplify the geometry calculation and to maintain consistent in direction relations, footprint can not be used as substitute for all features to compute direction relations. In order to determine the degree of generalization and the unknown slot values including *geometryType*, *rel\_Scale* and *topo\_Ref*, which will directly impact on the customization process of direction reference frame, an algorithm of extracting direction query characteristics is proposed. And the computation of direction relations can be divided into two types: footprint computation and entity computation.

This extraction algorithm of query characteristics (EAQC) divided mainly into five steps and the flow shows in Figure 2:

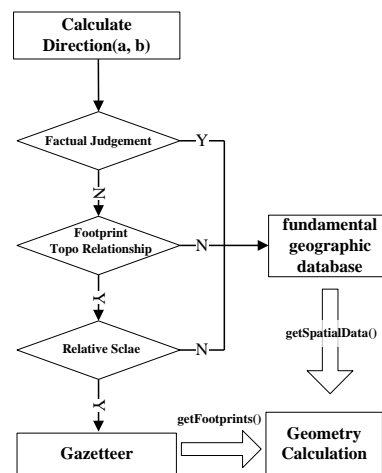


Figure 2 The extraction algorithm flow

#### Begin

Select two elements in geo-ontology:

{reference region: a, target region: b}

**Step 1:** To judge the facts. Execute a SPARQL query (Q1) to judge the explicit fact that contain(?a, ?b) or include(?a, ?b)

SPARQL: ASK {contain(?a, ?b) or include(?a, ?b)}

IF(Q1 is true)

*topo\_Ref*:= Interior;

goto Step 3;

**Step 2:** To calculate topological relations between footprint(a) and footprint(b).

IF(footprint(a) contain footprint(b))

*topo\_Ref*:= Interior;

ELSE IF(footprint(a) intersect footprint(b))

*topo\_Ref*:= Intersect;

goto Step 5;

ELSE IF(footprint(a) disjoint footprint(b))

*topo\_Ref*:= External;

**Step 3:** To compute relative scale between footprints using sampling grain model (Chen, 2011).

i) Computing Sampling Grain:

*rel\_Scale*(a):=*Gra*(a);

*rel\_Scale*(b):=*Gra*(b);

*Gra*=Max{ *Gra*(a), *Gra*(b)};

ii) Computing Standard Length:

$$l = \frac{\text{Max}\{\Delta X_A, \Delta Y_A, \Delta X_B, \Delta Y_B\}}{\text{Gra}} \quad \text{Gra} \neq 0 \quad (1)$$

where *Gra* = The number of sections which the longest side of MBR(A) and MBR(B) is divided into.

*l* = standard length

$\Delta X_A, \Delta Y_A, \Delta X_B, \Delta Y_B$  = the sides length of MBR(A) and MBR(B)

iii) Confirming the geometry type of objects:

IF (*Gra* equals to 1)

*geometryType(A)* := *point* AND

*geometryType(B)* := *point*;

ELSE IF ( $\text{Max}\{\Delta X_A, \Delta Y_A\}/l \leq 1$ )

*geometryType(A)* := *point* AND

*geometryType(B)* := *geometryType(footprint(B))*;

ELSE IF ( $\text{Max}\{\Delta X_B, \Delta Y_B\}/l \leq 1$ )

*geometryType(B)* := *point* AND

*geometryType(A)* := *geometryType(footprint(A))*;

ELSE

*geometryType(A)* := *Feature type(A)* AND

*geometryType(B)* := *Feature type(B)*;

goto Step 5;

Where *geometryType(footprint(A))* is the geometry type of footprint(A); *Feature type(A)* is the geometry type asserted in geo-ontology.

**Step 4:** To perform the geometry calculation from footprints.

**Step 5:** To perform the accurate geometry calculation from detailed shape of objects in spatial database.

**End**

### 3.3 Concretization of reference frame

After detail analysis and reasoning in semantic level, it is also significant to embody these query characteristics in physical level. Thus establishing the physical direction model should take topological relation, geometry type and constraint degree into account. However, existing direction models, such as cone-based, projection-based, MBR-based and so on, are suitable to different case. In this paper, Figure 3 shows several popular direction models as candidate models for geometric calculation, and do the analysis their applicability as follows:

- (I) Cone-based model. Mainly be suited for describing the direction relations between points.
- (II) Projection-based model. Mainly describing the direction relations that the geometry type of reference region is polyline, and the research region can be partitioned into to cardinal direction according to the distribution of geographical feature.
- (III) MBR-based model. Mainly describing geometry type of reference region is polygon.
- (IV) Interior MBR-based model. Mainly describing the direction that their topological relation is contain.
- (V) Sampling-based model or statistic-based model. Mainly describing the topological relation is intersect.

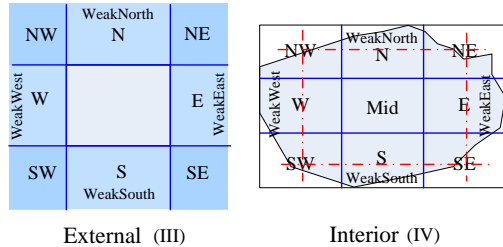
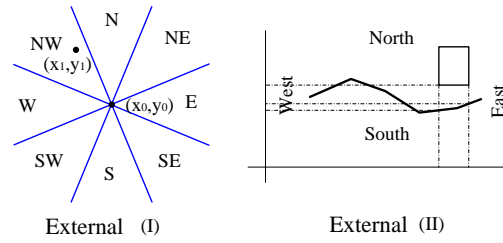


Figure 3. The main direction description models

The preferable direction relation model will be recommended to user according to different composition of query characteristics depicted in Table 1. Moreover, the uncertainty of geographical objects, the complexity of information processing and the inconsistency in the process of human cognition of spatial direction, all these factors can cause the vagueness of spatial direction relations. Thus, it is not enough to merely consider the extent of direction constrain described by user. Therefore, some fuzzy direction models are proposed for describing direction relations, which is not discussed here, and must define the fuzzy areas and the direction membership function.

	<i>Ref_Region Type</i>	<i>Tar_Region Type</i>	<i>Topo Reference</i>	<i>Reference Frame</i>
1	point	point	External	I
2	point	line		I
3	point	polygon		I
4	line	point		II
5	line	line		II
6	line	polygon		II
7	polygon	point		III
8	polygon	line		III
9	polygon	polygon		III
10	polygon	point	Interior	IV
11	polygon	line		IV
12	polygon	polygon		IV

Table 1. Composition of query characteristic

## 4. EXAMPLE FOR MULTI-MODE DIRECTION QUERY MODEL

There is a geo-ontology “GeonameOnto” to support the spatial reasoning and supply useful knowledge, which is built with Web Ontology Language (OWL) after analyzing the research and application of gazetteer. And the fundamental geographical data of China is used to supply the detail geometric information; the scale was 1:250,000. The footprints were acquired from Geonames webservice. The structure of direction relation query model is presented in Figure 4.

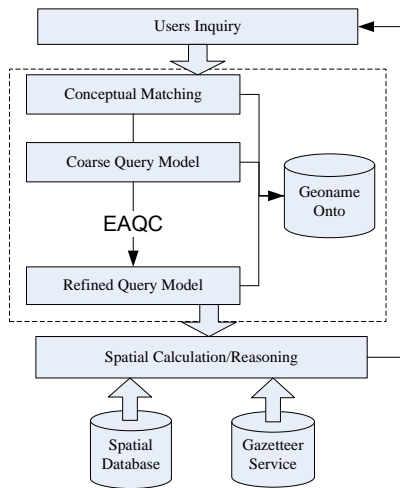


Figure 4. The structure of direction relation query model

Take Q1 for example,  $Q1 = \{ \text{“Provinces lie to the north of Changjiang river”} \}$ , a series complete framework instance can be produced with this approach. They were mainly divided into two categories: Footprints-based and Actual-objects-based.

Footprints-based, include two instances with different degree of generalization, is to compute direction relations based on the simplified shape or footprint. Their model instance is as follows:

**Instance 1** :  $\langle \langle \text{River, Line, } Gra(a) \rangle, \langle \text{Province, Point, } 1 \rangle, \langle 2, \text{External, II} \rangle, \langle \text{north, } 0.75 \rangle \rangle$

**Instance 2** :  $\langle \langle \text{River, Line, } Gra(a) \rangle, \langle \text{Province, Polygon, } Gra(b) \rangle, \langle 2, \text{External, II} \rangle, \langle \text{north, } 0.75 \rangle \rangle$

Actual-objects-based is to compute direction relations based on the real objects.

**Instance 3** :  $\langle \langle \text{River, Polygon, } Gra(a) \rangle, \langle \text{Province, Polygon, } Gra(b) \rangle, \langle 2, \text{External, V} \rangle, \langle \text{north, } 0.75 \rangle \rangle$

There are 33 geographical entities as the instance of Province (the second tier of administrative division) in GeonameOnto, where six candidate target regions to calculate direction relations with Instance 1, sixteen with Instance 2, and eleven with Instance 3. At last we get 17 provinces as the results of calculation, and the detail is shown as follows:

Instance	Provinces
$R_1$	Beijing, Heilongjiang, Jilin, Liaoning, Neimenggu, Ningxia, Shanxi, Shandong, Tianjin, Xinjiang
$R_2$	Hebei, Henan
$R_3$	Qinghai, Hubei, Jiangsu, Anhui, Jiangxi

Table 2. the detailed result of Q1

## 5. CONCLUSION AND FUTURE WORK

In this paper, we proposed a new conceptual model of spatial direction relations to exploit the implicit information and to achieve the semantic expansion of user query based on geographic ontology. User's intention can be restored by taking some influence factors into account to extend the traditional

triplet model. And this omitted information will be reappeared in conceptual model as query characteristics, which will be transmitted to the physical model. Then a complete instance can be obtained by using the extraction algorithm for query characteristics and some inference based on knowledge base, and it can recommend the most suitable direction calculation model according to actual case. At last, the validity and practical of this direction query model is verified by the experiment results. And it is shown that exploiting the implicit information in spatial direction relation query and making a reasonable understanding on the users intention can effectively not only reduce the computational cost, but also improve the query accuracy. As a future research, we will further consider to provide more accuracy measurement of relative distance from the perspective of geometry and geospatial semantics. Meanwhile, it is necessary for establishing a sound knowledge base about description models to make a more detailed analysis on the characteristics of existing models.

## 6. REFERENCE

- Carsten K, J. Krzysztof, et al. 2009. Mohamed An agenda for the next generation gazetteer: geographic information contribution and retrieval, Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, November 04-06, 2009, Seattle, Washington
- Chen, D. and C. H. Zhou, et al., 2011. A fuzzy decision method for direction relations based on spatial cognition, Geoinformatics, 19th International Conference on.
- Chen, T. and M. Schneider, et al., 2010. The Objects Interaction Matrix for Modeling Cardinal Directions in Spatial Databases, DASFAA 2010, Part I, LNCS 5981, pp. 218–232.
- Deng, M. and Z. Li., 2008. A statistical Model for Directional Relations Between Spatial Objects. Geoinformatica, 12(2), pp. 193–217.
- Du, S. H. and L. Guo, et al., 2008. A model for describing and deriving direction relations between overlapping and contained regions. Information Sciences, 178( 14), pp. 2928- 2949.
- Du, S. H. and L. Guo, et al., 2010. A scale-explicit model for checking directional consistency in multi-resolution spatial data, International Journal of Geographical Information Science, 24(3), pp. 465-485.
- Goyal, R. K. and M. J. Egenhofer 2000. Consistent queries over cardinal directions across different levels of detail. Database and Expert Systems Applications, Proceedings. 11th International Workshop on.
- Jin, X. and H. Y. Geng, et al, 2009. On Vagueness of Cardinal Direction Relationships in Difference Cognitive Scenarios. Acta Scientiarum Naturalium Universitatis Pekinensis, 45(6), pp. 97-104.
- L. Hill. 2000. Core elements of digital gazetteers: Placenames, categories, and footprints. In J. Borbinha and T. Baker, editors, Research and Advanced Technology for Digital Libraries - 4th European Conference, ECDL 2000 Lisbon, Portugal, September 18 - 20, page 280 – 290.

Mata, F. and F. Fonseca, et al. 2007. Geographic Information Retrieval by Topological, Geographical, and Conceptual Matching, GeoSpatial Semantics, Springer Berlin, Heidelberg. Vol. 4853, pp. 98-113.

Wang, Y. D. and J. J. Dai, et al, 2007. Geo-ontology Design and Its Logic Reasoning. Geoinformatics2007. Proceedings of SPIE, 0277-786X, V.6753.

#### **ACKNOWLEDGEMENTS**

This research was supported by “National Natural Science Foundation of China Grant 40971232”; “National Key Technology R&D Program of China Grant 2007AA120201”; “the Fundamental Research Funds for the Central Universities China 201161902020016”