Advanced Science and Technology Letters
Vol.45 (CCA 2014), pp.49-54
http://dx.doi.org/10.14257/astl.2014.45.10

Organization and Retrieval Method of Multimodal Point of Interest Data Based on Geo-ontology

Ying Xia, Shiyan Luo, Xu Zhang, Hae Yong Bae

Research Center of Spatial Information System, Chongqing University of Posts& Telecommunications, Chongqing, China
xiaying@cqupt.edu.cn

Abstract. With the rapid development of mobile Internet and wireless communication technology, mobile geographic information service (MGIS) is promising and applied in many fields. As the core content of MGIS, point of interest (POI) exhibits the properties of multi-source, heterogeneity and media convergence. Traditional spatial data organization method and semantic expression cannot meet the needs of multimodal point of interest (MPOI) data access in the mobile environment. In this paper, the Geo-ontology model is adopted to organize the POI data, it satisfies the requirement of multimodal fusion and semantic sharing in the mobile environment. On this basis, we propose a MPOI retrieval algorithm.

Keywords: Mobile Environment, Multimodal Point of Interest, Geo-ontology, Data Organization, Retrieval

1 Introduction

Point of interest (POI) is the core content of MGIS [1]. Its descriptive information becomes multi-source and heterogeneous gradually. The content of the POI data may come from different providers and usually combined by different media types, such as text, image, audio, and video. In this paper, we define it as multimodal point of interest (MPOI). To organize MPOI data reasonably and offer an effective retrieval in the mobile environment plays a fundamental role in driving the development of MGIS patterns and contents in the new period.

2 Related Work

POI is a fundamental concept in geographic information system (GIS). It refers to the geographic entity object which can be abstracted into a point. The spatial data refers to the date related to reference space location, which can express the status attribute of entities and process in the objective world [2]. Hence, POI belongs to the spatial data.

At present, there are mainly three kinds of methods on the organization of spatial data. They are topology-based data model, object-oriented data model and entity-relationship data model [3] [4] [5]. Among them, the entity-relationship data model is

ISSN: 2287-1233 ASTL

the most popular spatial data organization model. But it reveals the following weaknesses when it is used to organize MPOI data.

(1) No concern on the semantic information of POI data.

(2) POI is similar to the point feature (the features in GIS can be divided into point, line, and polygon). However, there are differences between POI and point feature [6], so conventional relation database cannot meet the requirement of MPOI storage.

(3) Simple organization method of relation data requires a large amount of storage space, which fails to meet the demand of mobile applications.

Ontology comes from philosophy. It is introduced to GIS as the concept of Geo-ontology. As a formalized, explicit specification of a shared conceptualization in GIS [7], Geo-ontology can be used to describe the detailed contents and hierarchical relations of geographical spatial concepts [8]. Because of its advantage on the expression of semantic information and spatial relationship, it is suitable for organizing the spatial data.

Therefore, we use Geo-ontology model to organize the multi-source and heterogeneous MPOI data in the mobile Internet. Then the requirement of multimodal data access and semantic sharing in the mobile environment can be met. On this basis, we propose a MPOI retrieval algorithm.

3 Geo-ontology Based MPOI Data Organizing Method

3.1 The Characteristics of the MPOI Data

Through our analysis of the mass MPOI data existing in the mobile Internet, we find that it has the features as followed:

(1) The source of data is multiple, and the amount of data is large.
(2) The MPOI data is multimodal, and the forms of MPOI data are abundant.
(3) The relation of different MPOI data is complex.
(4) The application of the MPOI data depends on the mobile environment.

3.2 Geo-ontology Model Building

Geo-ontology abstracts out the knowledge, information and data in GIS as an object or entity. Then they can form a system according to their relation. In this way, the Geo-ontology services for establishing the relation of GIS and geographic reality, also for sharing and organizing the geographic information [8]. According to the principles of building the Geo-ontology, we define the Geo-ontology as a quadruple model:

$$Geo\text{-}o = (C, R, I, A) \quad (1)$$

In (1), *Geo-o* represents the Geo-ontology, *C* indicates the set of concepts, *R* means the relation of the concepts or the relation of the instances (including the class inheritance relation, the spatial relation, and the semantic relation), *I* express the set of instance, and *A* denotes the set of inference axiom.

We use Upper Level Ontology design mode and Seven-Steps method to build the POI ontology in the protégé environment. The building of the POI ontology is according to Geo-ontology model, the characteristic of MPOI data, and the needs of the practical implication. We also reference the document from ESRI Company [9]. The structure of the ontology is shown in Fig. 1.

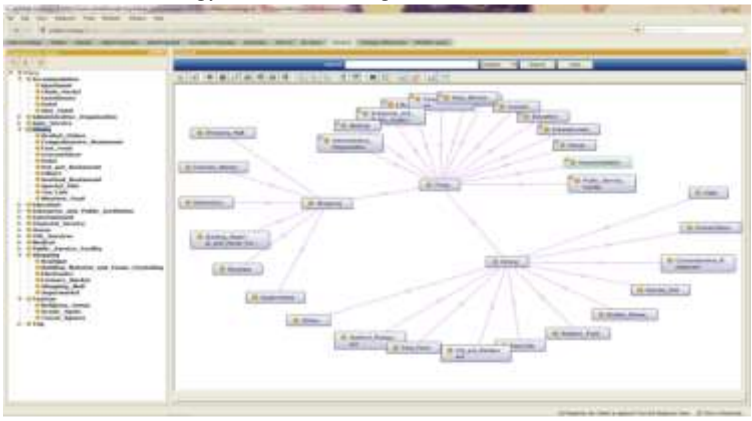

Fig. 1. Concept structure of POI ontology

3.3 The Relation of the Geo-ontology Entity and the MPOI Data

Multi-source heterogeneous MPOI data has semantic difference, which goes against the uniform organizing and sharing. However, the geographic entity, which is described by the Geo-ontology, has obvious hierarchical affiliation relation and plentiful semantic relation. The realized and conceptualized MPOI data can be abstracted out as an entity of a concept node in the Geo-ontology. The spatial relation of MPOI data also can be described by the Geo-ontology. Thus, we can uniformly describe and map the MPOI data in a reliable and standardized way. The Geo-ontology entities (which has semantic description, relation description, extended vocabulary, and one of the many types of media cluster center), can be mapped to the MPOI objects.

3.4 The OWL Description of the Instance Object

The description language of Geo-ontology should have a good-defined grammar, high-efficient inference support, formalized semanteme, sufficient ability of expression, and convenient presentation. OWL language, based on RDF and RDFS, is a sufficient ontology-expression and inferable language [10].

In this paper, the POI ontology entity describes the semantic information and the relation information with other POI. Take the “Jiafu Hotpot (Huang Jueya)” in the “hot-pot restaurant” category as an example. Its OWL description is shown in Fig. 2.

Fig. 2. The OWL description of instance object--"Jiafu Hotpot (Huang Jueya)"

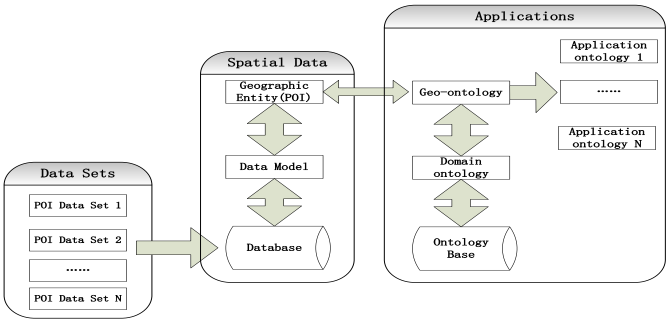

Fig. 3. MPOI data organization framework

3.5 Geo-ontology Based MPOI Data Organization Framework

The designed Geo-ontology based MPOI data organization framework is shown in Fig.3. In this framework, we collect MPOI data from Internet. After realizing, analyzing, abstracting, and conceptualizing, we map these data to the POI objects in the database. This kind of organizing method can conveniently manage the semanteme and the spatial relation of the multi-source heterogeneous MPOI data.

4 MPOI Retrieval Algorithm

4.1 Global Correlation

In practical, users usually need a comprehensive or personalized result, more than a simple result from the keyword matching or the position operations. So the correlation measurement result should be a comprehensive one which integrates the low-level, semantic and spatial features of POI. We define it as the global correlation δ.

$$\delta = \omega_1 \alpha + \omega_2 \beta + \omega_3 \gamma \qquad \textbf{(2)}$$

In (2), α represents the low-level feature correlation of POI, obtained from calculating the Euclidean distance between two eigenvectors. β denotes the semantic correlation of POI, acquired from calculating the hierarchical relation between two concepts or from the extended vocabulary. γ indicates the spatial correlation of POI, gained from calculating the spatial distance. ω_1 , ω_2 , ω_3 are weighting factors corresponding to the three correlations. Their values must conform to the formula $\omega_1+\omega_2+\omega_3=1$ and different values combination reflects users' preferences. The global correlation brings a more comprehensive, accurate and personalized result.

Table 1. The MPOI Retrieval Algorithm Description

MPOI Retrieval Algorithm
Input：Retrieval request *Q*
Output：Retrieval result set *T*
Step 1: Call the semantic acquirement algorithm to obtain the corresponding semantic set *S* of *Q*;
Step 2: Traverse all the concept nodes under the POI ontology. Then match them with *S* to get the concept node with the highest matching value as the candidate concept *C* of *Q*. If *C* is empty, *T* is null, go to the end;
Step 3: Judge the category of the *Q*, if it is a text , go to step 4; otherwise go to step 5;
Step 4: Conduct the global correlation measurement on all the instance objects in concept *C* with *S*. The semantic relation is measured by matching with the extended vocabulary, where the correlation coefficient of the low-level characteristics is set to 0. Then the instance objects, whose correlation coefficient is larger than the threshold ε_1, can be obtained in the instance object set *R* .Go to step 6;
Step 5: Conduct the global correlation measurement on all the instance objects' same media cluster center in *A* with *Q*. Then the instance objects, whose global correlation value is larger than the threshold ε_2, form the set *R*. Go to step 6;
Step 6: If *R* is empty, return null; otherwise return the corresponding POI data of *R* as *T*.

4.2 The Description and Analysis of the MPOI Retrieval Algorithm

The MPOI retrieval algorithm description is shown in Table 1. This algorithm supports multimodal retrieval requests and is able to realize personalized retrieval.

We assume the total number of MPOI data is n. Traditional retrieval algorithm based on the keywords has the time complexity as O(n).In our algorithm, the number of POI ontology underlying concept nodes is a. And the maximum number of instance objects in a concept node is m. Since a is much less than m, the time complexity of our algorithm is O(m). Obviously, m < n, thus O(m) < O(n). Although O(m) and O(n) have the same order of magnitude, the time complexity of our algorithm is still decreased compared to the traditional one. It dictates that our algorithm improves the efficiency of the whole retrieval process.

Our algorithm needs extra space overhead for saving the Geo-ontology data. However, the Geo-ontology model needs very little storage space. It can be ignored compared with the space overhead for saving the MPOI data. It means our algorithm sacrifice very little storage overhead to get an improvement of the retrieval efficiency. In addition, we save the Geo-ontology data in the mobile terminals, and the POI data in the servers. In this way, the data organization and translation are more suitable for the mobile environment.

5 Conclusion

Traditional organization method of spatial data cannot meet the needs of the MPOI data organization in the mobile environment. In this paper, we analyze the data characteristics of MPOI in the mobile environment, and adopt a Geo-ontology model to organize the POI data. On this basis, we propose a retrieval algorithm and correlation measurement of MPOI. We hope that all the work can provide certain references for the development of MGIS in mobile Internet.

Acknowledgments. This work is supported by National Natural Science Foundation of China (41201378), Natural Science Foundation Project of Chongqing CSTC (cstc2012jjA40014), and Doctoral Startup Foundation of Chongqing University of Posts and Telecommunications (A2012-34).

References

1. Lin, C.S., Chiang, M.F., Peng, W.C.: An Event-Based POI Service from Microblogs. In: 13th Asia-Pacific Network Operations and Management Symposium (APNOMS), pp. 1-4(2011)
2. Yu, F.Y.: Research and Application of Spatial Data Organization and Management. Dalian: Dalian University of Technology(2003)
3. Paul, M., Ghosh, S.K., Acharya, P.S.: Enterprise Geographic Information System (E-GIS): A Service-based Architecture for Geo-spatial Data. In: International Conference on Geoscience and Remote Sensing Symposium (IGARSS), pp. 229-232(2006)
4. Zhu, R., Qiu, M., Hu, Y.N.: Research on the Key Technolog of Multi-sources Data Integration for Spatial Data Updating. Engineering of Surveying and Mapping, 22(4): 22-25(2013)
5. Chen, H.X., Sun, Q., Xiao, Q., Xiao, J.H.: Application of Spatial Data Fusion in the Production and Updating of Spatial Data. Geomatics and Information Science of Wuhan University, 39(1): 117-122(2014)
6. Yu, Y.P.: Dynamic Generalization Methods of POI for the Representation of Mobile Map. Nanjing: Nanjing Normal University(2012)
7. Gao, B. Li, L.: Domestic Ontology-based Information Retrieval Research in Past Five Years. Library, 2: 57-59(2012)
8. Wang, Z.F., Lei, J.Q.: A Study of Spatial Information Sharing Model Based on Geological Hazard Domain Ontology. In: Fifth International Conference on Computational and Information Sciences (ICCIS), pp.268-270(2013)
9. POI Classification, Stratification and Attribute Structure Document of Geographic Information Public Service Platform. www.esrichina-bj.cn
10. Chen, Y.P., Wei, L.L.: Building of Domain Ontology about Tourism Based on OWL. Journal of Xi'an University of Posts and Telecommunications, 16(1): 78-82(2011)

