

Research on the Integration Method of Dynamic Heterogeneous Spatial Information Plotting Data

Xiao-lan Xie^{1,2}, Zhen Long³, Wen-qi Liao⁴ and Xiu-juan Guo¹

¹ College of Information Science and Engineering, Guilin University Of Technology, 541006, Guilin, Guangxi, China

² Guilin University of Technology, Guangxi Key Laboratory of Spatial Information and Geomatics, 541004, Guilin, China

³ College of Mechanical and Control Engineering, Guilin University Of Technology, 541006, Guilin, Guangxi, China

⁴ College of Civil Engineering and Architecture, Guilin University Of Technology, 541004, Guilin, Guangxi, China

Abstract. According to the dynamic heterogeneous spatial information plotting data, this paper presented a new integration method about dynamic heterogeneous spatial data, studied the spatial information plotting object and established the semantic ontology database. It declared the mapping rules in ontology, solved the semantic differences of spatial information plotting objects and finished the automatic or semi-automatic integration of plotting content to integrated automatic or semi-automatic of the content for different plotting situations (same name of different place, different name of same place etc) and different plotting data types (text, images, sound, video etc). So it has great significance to enhance the efficiency of the plotting data integration.

1 Introduction

Since 1990s, with the development of information and communication technology, as well as the wide use of remote sensing, geographic information system and satellite positioning technology, the acquisition mode of spatial information tends to be diversified, geographic spatial information is growing exponentially. Spatial information data show a distinct heterogeneity, that is, the characteristics of multi-source, multi-scale, multi temporal phase. However, due to the limitations of the data itself: a large number of spatial information data cannot be effectively used and the information needed by the user can not be satisfied in time, so the integration and update of spatial information data, information integration and fusion have become the cutting-edge tasks and key technology research in academic field and the key technology research. Due to the different application purposes of geographic data from different departments, it caused multi semantics, multi temporal and spatial, different storage format and the differences between data model and storage structure of spatial data in the same area and the same scale, it brings great difficulties to the data sharing and data integration between the GIS Department [1, 2, 3].

Especially in the different data production departments, in order to satisfy the different needs of users, the data standard and specification used are different. It exists great difference in the division of geographical elements, the classification of geographical elements, the definition of geographical elements, the definition of geographical element attribute, and the coding method of geographical feature, resulting the

diversity of multi-source heterogeneous geographic information in semantic, it is the main reason that the data could not be used directly with each other and share. Spatial information plotting data comes from different databases and complex relationship, so the data is not unified and not standardized, the relationships are confused.

Therefore, in view of the above mentioned problems about spatial information plotting data, it needs a new efficient integration method of heterogeneous spatial information [4, 5, 6].

2 Data Integration Method

In this paper, we study an effective integration method of spatial information plotting data, which is helpful for fully understanding the spatial information plotting data to solve the defects of the existing technology. The method mainly includes the following steps:

(1) The semantic ontology base construction of spatial information plotting object

It can study the problem of semantic dissension, determine the professional field covered by the semantic ontology of the plotting object and the purpose of applying ontology, list the important terms and concepts in the field, in order to build the ontology framework and design meta-ontology. It also encodes the domain ontology and verifies whether the ontology meets the requirements. All these have done to complete the construction of the semantic ontology base of spatial information plotting object. The local ontology is constructed according to the specific application, and

then the semantic mapping rules between heterogeneous spatial information are declared according to the relationship between ontology classes.

(2) The collection and extraction of the plotting data

① If there are business data of the front-end-processor, we can use the way of preposition exchange of the plotting data information to deal with the data. The department involved in the information exchange adopts the front-end-processor to achieve data exchange, which is composed of the operating system, the pre exchange plotting information database, the information exchange communication interface and the information exchange bridge. The relevant departments need to send the data exchanged to the front-end-processor through the Department of LAN. The front-end-processor transmits the data to the information exchange platform. The information exchange platform processes the data, and then sends the data to the front-end-processor in the network according to the business process. The business department can access the data needed through front-end-processor and the real-time and dynamic submission of the spatial information plotting data is realized.

② The collection and extraction tool setting of the plotting data could realize the secure, reliable, stable and efficient the plotting data collection and extraction system between business data in different business database. The system takes some measures like collection, extraction and loading to convert the business data provided by some unit into the original database of the plotting data through the data transmission, information encryption, format conversion and system management.

(3) The analysis and processing of plotting data

The data filtering conversion system is to filter and transform the data that has been collected in the intermediate database. It combines with the semantic ontology database of plotting object and mapping rules in ontology, realizes the processing of different data and suspicious data according to a certain filtering and transformation rule. The system includes fault-tolerant management and spatial topological relation integration, saves the correct data filtered into the final database of plotting data and provides support for the construction of each data subsystem.

(4) Form the final database of plotting data

The final database is the final plotting database which confirmed after the former collection, extraction, analysis and processing of plotting data[7]. It provides support for other business system. Finally it realizes the classification management of data, authority management, security authentication, security audit by setting the safety management mechanism. It encrypts data in different security levels and realizes data backup.

3 THE EXPERIMENTAL METHOD

As is shown in Figure 1, according to the integration method of dynamic heterogeneous spatial data, the whole technical scheme diagram can be divided into four parts: establishing the plotting object semantic ontology database, collecting the plotting data, processing plotting

data, building the final database, it mainly includes the following steps.

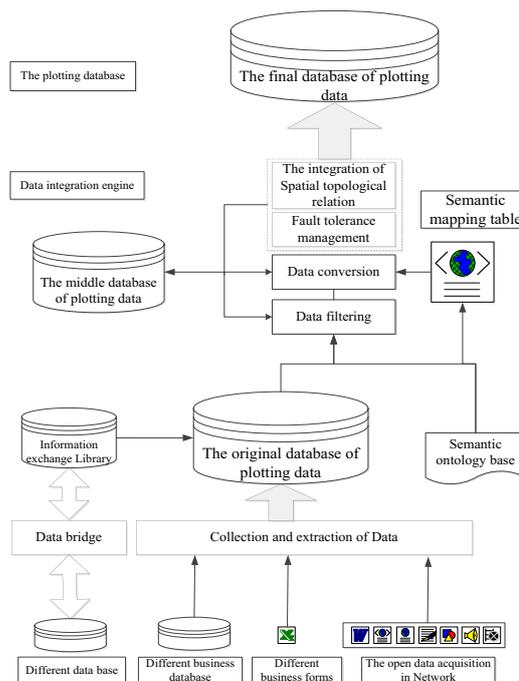


Figure 1. A flow chart of Integration method of dynamic heterogeneous spatial information plotting data.

3.1 Establishing the plotting object semantic ontology database

We use the open source Protégé ontology development tool developed by Stanford medical Informatics from Stanford University to express spatial information plotting object semantic ontology. Protégé ontology development structure displays in a hierarchy tree structure, it can easily add or edit classes, subclasses, attributes, instances, etc. It provides basic functions of ontology construction and achieves visual expression of ontology through the OWL Viz Tab. It describes geographic knowledge according to the form of class and attribute. Using ontology expression language based on OWL to clearly express the semantics and mutual relations of geographic concepts. It achieved by the use of open source Jena development kit which used to develop the semantic ontology application program, support the simple reasoning based on rules and implement mapping rules in ontology.

3.2 Collecting the plotting data

The system takes some measures like collection, extraction and loading to convert the business data provided by some unit into the original database of the plotting data through the data transmission, information encryption, format conversion and system management.

3.3 Processing the plotting data

3.3.1 Data processing method

For these plotting data from various sources, we can plot the object semantic ontology database by combining with the space information, then eliminate duplication, remove ambiguity, number, classify, nominate and so on, finally form the standardized data. According to the underlying demand of plotting data, after the data processing and filtering, we should combine with the declared semantic mapping rules, convert the original database data into the form that target data required. When analyze and process the data, we convert these data at the same time. Moreover, we make logical operation to the filtered data according to the requirements of the application, convert into the form that target data required, so that ensure the data validity. Then develop the fault tolerance rules in the case of ensuring the integrity of the data. Meanwhile, arrange the topological relation existed in the data, eliminate the errors in the data, ensure the quality of the data. According to the features of the object with different plotting data types (text, images, audio, video, etc.), we set up related data filtering rule, transformation rule, fault-tolerant rule and the topological relation collation rule, and take some related settings on the processing system, so as to improve the automation and efficiency of plotting data conversion and filtering.

The data processing method has the following characteristics:

①Use the platform architecture and design based on J2EE, use the current advanced struts, hibernate and other technologies to realize, and support Tomcat, Weblogic, WebSphere, Apusic and other kinds of application environments;

② The database based on various domestic and international mainstream are provided, such as the interface definition and information processing of Oracle, SQL server, Access, Sybase and other data systems. The information source of data processing could be defined freely, and support various operating system platforms such as UNIX, Windows and Linux;

③Support a variety of information pretreatment and management strategies of data processing, provide an unified, efficient and multi-tasking processing engine on the data processing, it can define and configure specific data processing and processing mode flexibly and dynamically;

④ Use the multi-thread processing mechanism, support a set of multi-tasking management functions to process systems, a variety of tasks can be defined freely and centralized management, and the operation interface and the use is very simple and clear;

⑤ Support the information processing based on a variety of file type database systems, such as SHP, DBF, Excel, XML, GML, etc. The processing strategies and configuration of information use the unified and simple management platform and tools;

⑥ Not only support automatic processing, manual processing and other various processing methods to the resource data, but also realize output, statistic, analyze and other further service functions to the processed data;

⑦Provide the service of information processing and processing log in details, it can define the requisite log category and detailed grade or degree by itself;

⑧ Provide fully independent security access and control system, each operation and resource use can be accomplished by defining roles and users;

⑨Use the B/S architecture design and development, using and managing the interface will be intuitive and simple, it does not need special training to users, the products provide the function of on-line use help at the same time.

3.3.2 Data processing flow

①First we should finish the data preprocessing on the imputation plotting data, mainly include taking the standardized treatment of the content of plotting data, for example, standardize some key fields whether include "-", standardize full width, half width and spacing of Chinese and English name;

②The extract of information center system need to take some automatic processing to the processed data. According to the certain rules and standards, we can plot the data by the data collection tools, after data collection and extract, form the original library of plotting data. Then filter the data through the data integration engine in accordance with a certain data filtering rule, so that form a filtered plotting data repository[8, 9, 10]. Because of the complex relationship between data sources, it is necessary to use a data transformation rule to the filtered plotting data repository for data conversion. In order to ensure the integrity of the data, we need to set the fault-tolerant rule, take some fault-tolerant managements (save, eliminate, etc) to the incorrect data. Meanwhile, integrate existed topological relations in the plotting data, eliminate topology errors, then form the final library of plotting data. At last, the final library will be transformed into the subject library of plotting data which provide the application to other business systems;

③In the data processing system, we could configure data processing methods according to a certain rule, then implement the data processing, so that these processed data can be collected into the final library of plotting data .

The flow chart of data processing is shown in Figure 2.

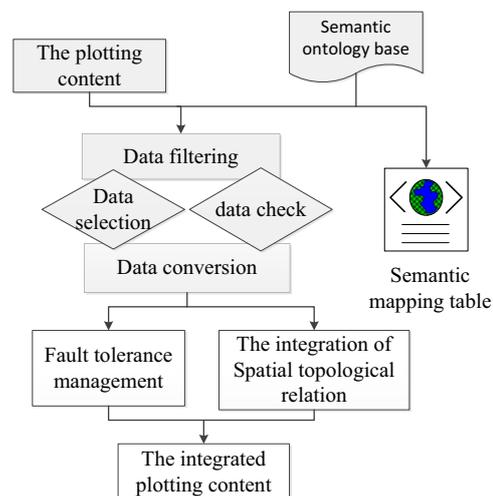


Figure 2. Flow chart of data content integration

3.4 Building the final database

The plotting data final database is the ultimate resource database which is formed by the data collection, extraction and processing analysis. It provides data support for other business system, and the classification of data management, authority management, security authentication, and security audit have been realized by setting the security management mechanism. Finally it encrypts the data in different security levels and realizes the secure backup of the data.

It mainly has the following advantages by taking the above overall frame structure:

① It uses the exchange front-end-processor and pre exchange of the plotting data resource information database, it can effectively separate switch network from department business network to ensure the safety performance of department business system and the plotting database;

② It uses the information exchange bridge to realize the on-line real-time exchange between the department business information database and the pre exchange plotting data information database, and meet the goal of online real-time plotting data information exchange between departments;

③ Establishing the information exchange platform with central exchange management function, which is convenient for system management and system expansion;

④ The system which is based on the information exchange platform can support all kinds of heterogeneous platforms or data sources, and facilitate the rapid expansion of the platform, in the premise of ensuring reliable, efficient and safe transmission of information;

⑤ The whole system has good expansibility, which is in line with the standard of national government information resource exchange system and other related standards;

⑥ By establishing the data integration mechanism of standard drawing object semantic ontology library, to constraint ontology through the terminology established, relying on the artificial intervention, establishing plotting object semantic mapping, so the spatial information semantic points difference can effectively be solved, and better data integration effect on painting has been realized.

4 Conclusion

As the different sources of the spatial information plotting data and the complex relationship, it leads to some features of the data, for example, the data is not unified, not standardized and the correlation between data is chaos. Therefore, aiming at the problem of the spatial information plotting data, this paper proposes a new method for the integration of dynamic heterogeneous spatial information data. According to the experiment results, the integrated method can increase the efficiency of data integration, so it has certain practical significance.

Acknowledgement

This research work was supported by the National Natural Science Foundation of China (Grant No.61540054), National High Technology Research and Development Program 863 under Grant No. 2013AA12A402, Natural Science Foundation of Guangxi Provincial under Grant No. 2013GXNSFAA019349.

References

1. Lin Fengcheng, Chung Lankun, Ku Wenyuan, Chu Linru, Chou Tienyin, Service component architecture for geographic information system in cloud computing infrastructure, Proceedings - International Conference on Advanced Information Networking and Applications, 368-373 (2013)
2. Yan Zhihui, Sun Qiang, Discussion on cloud GIS, WIT Transactions on Information and Communication Technologies, **57**, 387-394 (2014)
3. Diao Zhijian, Guo Song, A research of GIS software application based on cloud computing, Applied Mechanics and Materials, **513-517**, 2107-2110 (2014)
4. Schockaert Steven, Smart Philip D, Twaroch Florian A, Generating approximate region boundaries from heterogeneous spatial information: An evolutionary approach, Information Sciences, **181**, 257-283 (2011)
5. Li Weidong, Peng Shumin, Zhou Jinyan, Integration framework of multi-source heterogeneous spatial information and remote computing based on web services, 2010 2nd International Conference on Industrial and Information Systems, **2**, 457-460 (2010)
6. Wu Hao, He Chaoying, Liao Anping, Peng Shu, A framework for integrating heterogeneous spatial information and applications adaptively based on multi-agent and web service, Proceedings - 3rd International Conference on Multimedia Information Networking and Security, 253-257 (2011)
7. Condit Richard, Lao Suzanne, Singh Anudeep, Esufali Shameema; Dolins Steven, Data and database standards for permanent forest plots in a global network, Forest Ecology and Management, **316**, 21-31 (2014)
8. Bergamaschi Sonia, Guerra Francesco, Orsini Mirko, Sartori Claudio, A new type of metadata for querying data integration systems, SEBD 2007 - Proceedings of the 15th Italian Symposium on Advanced Database Systems, 266-273 (2007)
9. Wang Xin, Huang Linpeng, Xu Xiaohui, Zhang Y.I., Chen Junqing, Journal of Information Science and Engineering, **27**, 681-695 (2011)
10. Wang Xiaobin, Wang Wangx, Bao Mingyu, Research into framework of real-time data integration system, Applied Mechanics and Materials, **310**, 605-608 (2013)