

Spatial Data Integration using Ontology-Based Approach

S. Hasani^{a,*}, A. Sadeghi-Niaraki^a, M. Jelokhani-Niaraki^b

^a Department of GIS, Faculty of Geodesy and Geomatics Engineering, K.N.Toosi University of Technology, Tehran, Iran -
saeed.hasani01@gmail.com, a_sadeqi313@yahoo.com

^b Department of GIS & RS, Faculty of Geography, University of Tehran, Tehran, Iran - jelokhani83@gmail.com

KEY WORDS: SDI, Ontology, Ubi GIS, Data integration, Semantic Web

ABSTRACT:

In today's world, the necessity for spatial data for various organizations is becoming so crucial that many of these organizations have begun to produce spatial data for that purpose. In some circumstances, the need to obtain real time integrated data requires sustainable mechanism to process real-time integration. Case in point, the disaster management situations that requires obtaining real time data from various sources of information. One of the problematic challenges in the mentioned situation is the high degree of heterogeneity between different organizations data. To solve this issue, we introduce an ontology-based method to provide sharing and integration capabilities for the existing databases. In addition to resolving semantic heterogeneity, better access to information is also provided by our proposed method. Our approach is consisted of three steps, the first step is identification of the object in a relational database, then the semantic relationships between them are modelled and subsequently, the ontology of each database is created. In a second step, the relative ontology will be inserted into the database and the relationship of each class of ontology will be inserted into the new created column in database tables. Last step is consisted of a platform based on service-oriented architecture, which allows integration of data. This is done by using the concept of ontology mapping. The proposed approach, in addition to being fast and low cost, makes the process of data integration easy and the data remains unchanged and thus takes advantage of the legacy application provided.

1. INTRODUCTION

In today's world, the need for spatial data for various organizations felt so, many of the organizations have begun to produce spatial data. In order to spatial data seamless access comply with security, these organizations needs an infrastructure called SDI (spatial data infrastructure).

Nowadays GIS trends are toward the Ubi GIS that it provide a platform which spatial data capabilities is available to anyone in any place, time and circumstances in a user-centric manner. Therefore, with the development of U-SDI (Ubiquitous SDI), SDI introduced to solve the challenges of Ubi GIS comply with access levels and security. Data Integration is one of the main problems in SDI and U-SDI, It seeks to integrate different data with different data sources in a way that's integrated data remains consistent and homogeneous.

Since some circumstances needs to obtain real time integrated data, sustainable mechanism is needed to process real-time integration. One example for the described circumstances is disaster management because it needs to obtain real time data from various sources of information, on the other hand, one of the challenges in described circumstances is the high degree of heterogeneity between different organizations data.

Various solutions are proposed for the Data integration problem, the simplest solution is integration of all data in a seamless database that is not possible for several reasons, first because of the complexity of security mechanisms which do this for several organizations, Secondly, this process is extremely costly and time-consuming. There are some other reasons such as complexity of the system maintenance and impossibility of using legacy applications. One of the proper way to solve this problem is to use the potential of the Semantic Web to real time integrate data. The Semantic Web, based on RDF and Ontology provides the possibility of detection and integration of data in proper way. This approach not only provides data security in organizations, but it reduced costs and consume less time as well as the possibility of using the results of the legacy applications.

According to semantic web capabilities, it can be used in SDI to solving some problems. Ontologies and semantic web technology can be used for data searching, data integration, or using ontologies for metadata in Spatial data infrastructure. Using semantic web improve data sharing and interoperability in SDI.

We use ontology as basis of our proposed method to provide sharing and integration capabilities for the existing databases. In addition to resolving semantic heterogeneity, better access to information is also provided by our proposed method.

In order to test the efficiency of our proposed method, we define a project in Tehran Electricity Company and resolved the aforementioned problems for that. our approach provides a platform to share information, notification of the available spatial and descriptive data and it offers the most appropriate way to integrate data from different electricity industry sources; to reduce parallel work and create an environment in which, different parts of the electricity industry can also coordinate with one another and achieves to their goals by the use of different technologies.

2. RELATED WORK

Integrating and sharing data among different resources and systems to achieve better and more efficient services are of significant importance. Many studies have been developed in this context, among which are following researches

SDI devises a collaborative environment to manage, store, and exchange spatial information (Groot et al., 2000). This infrastructure is usually considered as the fundamental infrastructure in the format of politics, standards, and access network which facilitates the sharing data among people (in organizations), governments, and countries (Vaccari et al., 2009)

Mohammadi in 2008 investigates in recognizing the potentials of SDI to efficiently integrate data, to design and develop key tools to integrate spatial data from different resources. To achieve this goal, a number of international and national Australian SDI projects are investigated as the case studies to

integrate spatial data. Divers problems and technical as well as non-technical challenges regarding spatial data integration are discussed. In this research, data integration, and validation tools, data integration model, as well as documents containing characteristics of metadata and data are presented (Mohammadi et al., 2009).

Rajabifard et.al in 2008 emphasize the need for development of a uniform SDI as a powerful platform to increase the efficiency and effectiveness of data management among the vast majority of society, who are not spatially aware. In addition, they attempt to determine and identify the properties and basic indicators to use the uniform SDI model, and existed obstacles in operating this model. The uniform SDI enhanced sharing of data and cause a better connection among organizations which results a better decision. This research is presented to facilitate development of an extended framework to support areas of terrestrial and marine applications (Rajabifard et al., 2008).

Manso et al in 2009 provided an integrated model that facilitates exchange of data services and application on SDI. This model defines interoperability levels for spatial data infrastructure (no interoperability, Technical interoperability, Syntactic interoperability, Semantic interoperability, Pragmatic interoperability, Dynamic interoperability, Conceptual interoperability). This paper show the role of metadata elements for reach to different interoperability. This model can help developers for implementation of spatial data infrastructure and provide better services among interoperable system.

Gimenez et al in 2013 offered a viewpoint to create spatial ontology to present spatial objects in national SDI of Brazil. This point of view provide detection and integration of information. Proposal spatial ontology offer a basis to integrate several semantic heterogeneous sources in SDI. This ontology contain of three sub-ontology for spatial names ontology, metadata profile ontology and spatial entities ontology based on spatial objects. (Gimenez et al., 2013)

Janowicz et al in 2010 offered a Semantic Enablement Layer for SDI. This layer provided integration of semantic services; and solved semantic interoperability creation problems like creation of ability of services in data exchanging in a meaningful and minimum human manipulation method. The proposal model consist of two group; there is one group to manage and access ontologies and the other group is to present conclusion ability in SDI. (Janowicz et al., 2010)

Cruz et al in 2009 present an ontology-based method for integrate heterogenetic data across different sources. It involves an automatic ontology alignment process and algorithm for query among the distributed data sources.

3. METHODOLOGY

Various organizations' existing databases that have developed with the specific purpose are used only by the organization software. Objective data producing for each software is one of the causes of semantic heterogeneity of data as well.

The current data is expensively produced under the relational databases so this kinds of data cannot be abandoned and produce OWL or RDF data. Our proposed solution is linking the ontology with existing relational database then we can use ontologies as to better data sharing and various data integrating and also our proposed method in addition to system capabilities increasing, allows more complex queries to the databases which it can't be implemented by SQL language. These factors can help data integration under the SDI platform.

The proposed method (as depicted in Fig. 1), is done in three steps (as depicted in Fig. 2), the first step is identification of the object in a relational database, after the object is identified, the

semantic relationships between them are modelled then the ontology of each database is created.

Ontology of database with consideration of entities within it and the relationship between them in the real world is designed. At this point, part of the knowledge in the real world is modelled by the ontology.

Then, in a second step the relative ontology inserted into the database and the relationship of each class of ontology inserted into the new created column in database tables. After the second step performed in each database, objects and relations between them can be achieved by querying on ontology. This makes the queries that are not answerable, easily turned to answerable queries by the defined ontology model. In order to have the proper relation between the databases, and a web service is created for each integration scenario (Part A and B in Fig.1)

Proper ontologically modelling of the relationships between objects in the real world and existing relational databases make it possible to add the Semantic Web ontology capabilities to existing data, without any changes in the database and spending high cost to modify the database' connected software and generating new data.

In the next step, a platform based on service-oriented architecture allows integration of data. This is done by using the concept of Ontology mapping. The ontology obtained from the pervious process are related to each other by Ontology Mapping, This is done by the definition of ontology which is contained the different relations between existing ontologies. After this process, the possibility of access to data in a database is provided, so that the modelled relation makes the outputs homogeneous and consistent. And finally, the accessibility of users to services provided under the web services standards (Part C in Fig.1)

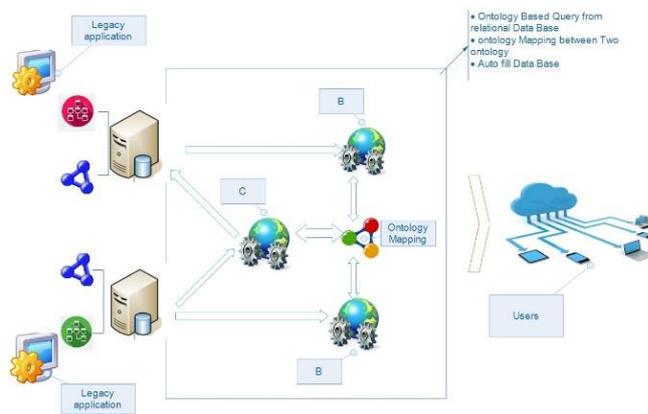


Figure 1. System Architecture

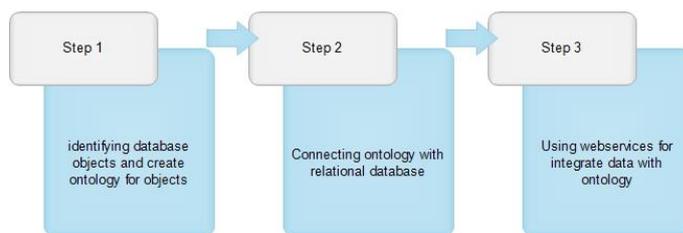


Figure 2. Steps for implementing method

development of Ontology based software and services. In addition, a sample of questions and answers based on Ontology also provided that cannot already been achieved by the SQL language.

The following scenario explained as an example, the X incident happened on an equipment. In the current situation, the maintenance team do not have much information. These people doing support actions only on the basis of experience. In these cases, in addition to the slow pace of support actions, some errors occur due to lack of information. As well appropriate incident management is not possible. The main need of disaster management in these situations is accessing to aggregated data. The next problem is to update the operational database after incident. Part of the information which used to update the operational GIS database is available, this work is now done manually, which is extremely time consuming and in addition it can be a source of inconsistency to the databases. Using the proposed method provide to users the ability to solve problems. When the incident occurs, the necessary information for the support of different databases integrated and in the form of Web services are available. Then, according to the circumstances the necessary changes on the related data database applied automatically and the database will be updated.

5. CONCLUSION

There are several ways to integrate data that each of which have their own characteristics. One of these methods that has been proposed in this paper uses ontology-based techniques to eliminate heterogeneity in the integration process. The proposed approach in addition to being fast and low cost, makes the process of data integration, data remains unchanged and thus take advantage of the legacy application is provided. Another purpose of this method is provide Consistency in multiple databases it means system can update same data in different databases and produce meaningful and consistent result. Seamless data updating between two databases reduce the cost of producing data and make data reliable for any user. In addition, the proposed method prevents data replication and its secondary problems. As described above combining the concepts of service-oriented architecture helps system to be available, flexible and more practical.

For future direction of research in this subject is comparing the proposed method with other methods for data integration. This method can be compared with ontology based methods or other approaches don't use ontology. In comparing can focus on some factors like time consuming, cost, efficiently, performance etc. another future work is study about complexity and efficiently of this method when it use it for high numbers of databases and show how ontology mapping and web services become more complicated.

REFERENCES

Cruz, L., and Xiao, H., (2008), Data Integration for Querying Geospatial Sources, Geospatial Services and Applications for the Internet, pp 110-134.

Gimenez, P. J. A., Tanaka, A., Baião F., (2013) A geo-ontology to support the semantic integration of geoinformation from the National Spatial Data Infrastructure, Proceedings of XIV GEOINFO, November 24-27, 2013, Campos do Jordão, Brazil.

Groot, Richard, and John D. McLaughlin, eds. Geospatial data infrastructure: concepts, cases, and good practice. Oxford: Oxford university press, 2000.

Janowicz, K., Schade, S., Bröring, A., Keßler, C., Maué, P., (2010), Semantic Enablement for Spatial Data Infrastructures, Transactions in GIS Volume 14, Issue 2, pages 111–129

Manso, M. A., Wachowicz, M., and Bernabé M., A., (2009). Towards an Integrated Model of Interoperability for Spatial Data Infrastructures, Volume 13, Issue 1, pages 43–67.

Mohammadi, Hossein. The Integration of multi-source spatial datasets in the context of SDI Initiatives. Diss. University of Melbourne, 2008.

Mohammadi, H., Rajabifard, A. and Williamson, I. P. (2009) Enabling Spatial Data Sharing through Multi-source Spatial Data Integration, GSDI 11 World Conference, 25-29 February 2008.

Vaccari, Lorenzino, Pavel Shvaiko, and Maurizio Marchese. "A geo-service semantic integration in spatial data infrastructures." International Journal of Spatial Data Infrastructures Research 4 (2009): 24-51.