A Proposed Architecture for Distributed and Version-Based Geospatial Data Sharing

Saeed NADI, Iran

Key words: Spatial Data Sharing, Versioning, Distributed Spatial Database, Spatial Data Infrastructure

SUMMARY

Developing systems to facilitate spatial data sharing have proven to be crucial for realizing e-government. A variety of organizations collaborate in land administration due to the diversity of activities in this business. In order to be able to adapt this various activities, it is required to enable organizations to share their data by utilizing appropriate frameworks and technologies. These frameworks and technologies assist the organizations in implementing seamless land administration. One of the most important parts of this framework is the facilities to share spatial data between these organizations, including legal rules and technical tools as a spatial data infrastructure. In this paper, a new spatial database architecture is proposed to utilize, update and analyze spatial data in a shared environment. The proposed model supports versioning of data and provides two level of quality control when an update introduced to data. Each update will firstly be applied on an isolated version named business version where will be checked to be acceptable regarding the business rules. Then accepted updates will be introduced to the second isolated version named technical version where will be checked to be spatially, topologically and cartographically true. After these quality controls the update will be released to be accessible by all the users. This ensures the accuracy, integrity and consistency of spatial database during distributed update process. A prototype system is developed to further investigate the proposed model. Distributed spatial database architecture is utilized for the development of the proposed model as the prototype system to satisfy data sharing in multiunit organizations like municipalities. The implemented system is used and tested as an infrastructure in the Shahinshahr municipality near Isfahan metropolitan in Iran using more than 100 users over 200 spatially referenced layers. The paper finally explains that service oriented architecture of the system can resolve many issues relating to the shared utilization, updating and analyzing spatial data.

A Proposed Architecture for Distributed and Version-Based Geospatial Data Sharing

Saeed NADI, Iran

1. INTRODUCTION

The growing quantity and quality of spatial data, which stems from a widespread need of different organizations for employing such information on their different levels of decision making has led the local governments to spend a huge amount of resources invested in the production of this information. In this regards a wide variety of spatial information are produced and utilized in parallel (Masser, Rajabifard, & Williamson, 2008).
Developing a framework to share this data between different sectors make it possible for one section to use information produced by other sections. This will also have the benefit of online, distributed and incremental spatial data updating which is a challenging task todays (Donaubauer , 2005). Moreover sharing spatial data may leads to save money and avoid of duplication in data production activities.
Sharing of spatial data involves more than simple data exchange. In order to facilitate the spatial data sharing, spatial stakeholders require dealing with many issues including data integration, representation and updating (Onsrud & Rushton, 1995).
Before the 90s, organizations purchased geographic information systems with a native, spatial data model. The non-relational file structures of these systems were optimized for fast access to data, were relatively easy to distribute between different sites the offline. However, the ability to share data online among users within an organization was limited. Data sharing between organizations with different GIS vendor systems was limited to data converters, transfer standards, and later open file formats. Sharing spatial data with other core business applications was rarely achieved (ESRI, 2003). Gradually, GIS models evolved into Geo-relational structures, where related attribute data could be stored in a relational database that was linked to the file-based spatial features. In this model supporting large data layers required the use of complex tiling structures to maintain performance, and sharing spatial information with other core business applications was still not possible. In the 90s, new technologies emerged that enabled spatial data to be stored in relational databases supporting large, non-tiled, continuous data layers. These Geo-relational databases (also referred to as spatial databases) could be embedded within core business applications where the sharing of spatial features became possible (Mohammadi, Rajabifard, & Williamson, 2009).
In this paper a spatial data sharing framework is proposed which facilitates online updating in a distributed environment using spatial database systems. The architecture is based on version concept which promises the quality control of the data during updating processes. The framework then is implemented and tested in the municipality organization of Shahinshahr, a city near Isfahan metropolitan in Iran. Utilization of this framework in Shahinshahr and sharing spatial data and updating between different sectors shows the efficiency of the proposed architecture.
Following, some technical aspects of spatial data updating is described in section 2. Section 3, describes the proposed architecture and in section 4 the implemented systems are illustrated. Finally, conclusions and remarks are described in section 5.

2. SPATIAL DATA UPDATING METHODOLOGY

One of the most challenging issues in multi-unit organizations is to provide updated spatial information between their different sections. There are often two methods used. In the first method changes applied to spatial information in specific times by a technical division. In the second method, changes applied by users during their daily activities named "incremental and shared spatial data updating". The most advantage of this method is that the most updated data are available in real-time.

2.1 Incremental and Shared Spatial Data Updating

There are four steps for incremental and shared data updating:

1- Migrating from file based maps to spatial databases
2- Distribute spatial information between different divisions
3- Provide spatial data updating tools to users
4- Provide quality control mechanisms

In file based map usage, each user has a distinct version of data and is not aware of the changes applied by other users unless they get a copy. Spatial databases provide a suitable platform to distribute maps between different users. In spatial database systems different users access to same map from different locations using different access level. Some users have editing permissions to some layers which are accessible for just viewing by other users. Updating tools play a very important role in this method. Users should use the same tools as they worked with in the file based method. This means that the architecture must let CAD software to be used for editing. As data are accessible in a shared environment, the edits by one user must be accepted by others. This makes the quality control very important steps.

2.2 Distribute Spatial Database Systems

Spatial database systems are special database which is optimized and structured to store and query geometric data such as point, line and polygon in the same time with their attributes. There exist approaches to organize and structure spatial data called database model. One of the most common models uses tables constructed from rows and columns named relational database models.

In spatial relational databases in addition to typical SQL queries there are a variety of special operations such as spatial measurements for computing length, area and etc., special functions to modify existing features, spatial predicates to define relationships between features, geometry constructions to create new geometries, spatial query to retrieve information and finally spatial indexing methods to speed up database operations.

2.3 Versioning

Versioning is a mechanism that enables concurrent multiuser spatial data editing in spatial database systems. It uses a concurrency data-locking model, which means no locks are applied to features and rows during the editing process. This mechanism provides support for many users creating and maintaining large amounts of GIS data in a central location. In many cases, multiple users need to edit the same data at the same time. In other words, they require

concurrent multiuser geodatabase editing.
Versioning manages and records states of individual features as they are edited. It is the basis for multiple users accessing and editing data simultaneously. Conceptually, a version represents an alternative, independent and persistent view of the database. It supports multiple concurrent editors and does not duplicate the data. A version references a specific state of the database. It contains all the datasets and evolves over time.
Versioning make it possible to distribute spatial data between different divisions and at the same time let us to control the quality and the correctness of the edits applied by different users. The last one can be done by comparing versions related to different users.

3. DESIGNING A DISTRIBUTED ARCHITECTURE FOR VERSION-BASED GEOSPATIAL DATA SHARING

Figure 1 illustrates general schema proposed for sharing spatial data between different divisions.

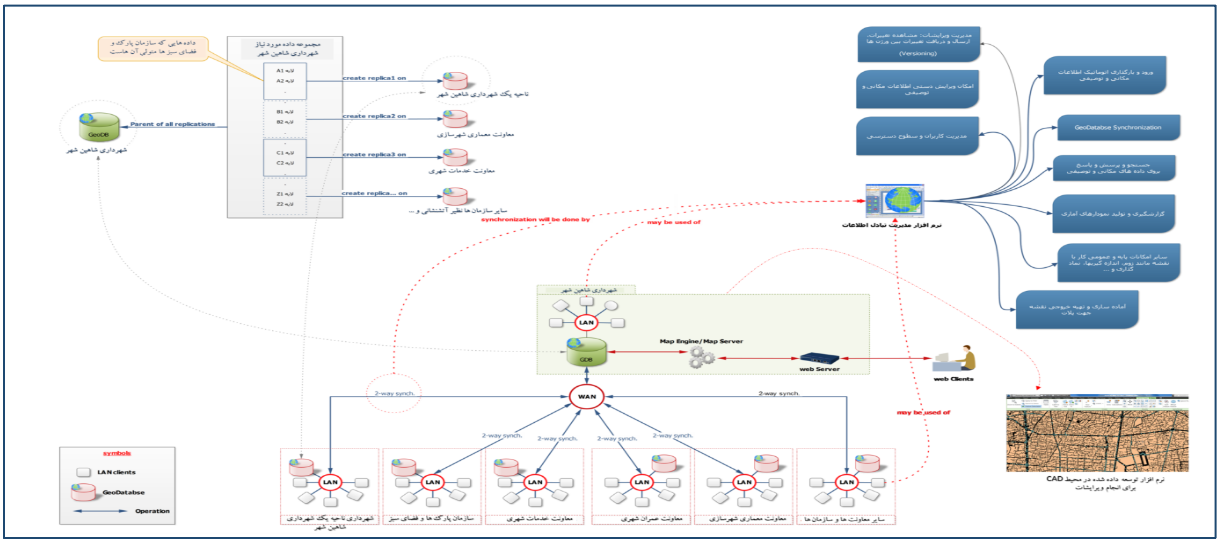

Figure 1. General proposed architecture

As illustrated in figure 1, a central database is used to store spatial data seamlessly. There are a number of distributed databases replicated from this central database. These databases are located in different divisions in which users connected to and edit, observe or analyses related data. Each user has his/her own version. Any modifications applied to data from each user just affect the related version. Two levels of quality control are designed to assure the validity and the correctness of modifications. The first level is business level, in which edits are investigated to be correct from a business point of view. The second level is technical level in which all edits are checked to assure that are geometrically and logically true. Figure 2 illustrates proposed versioning schema.

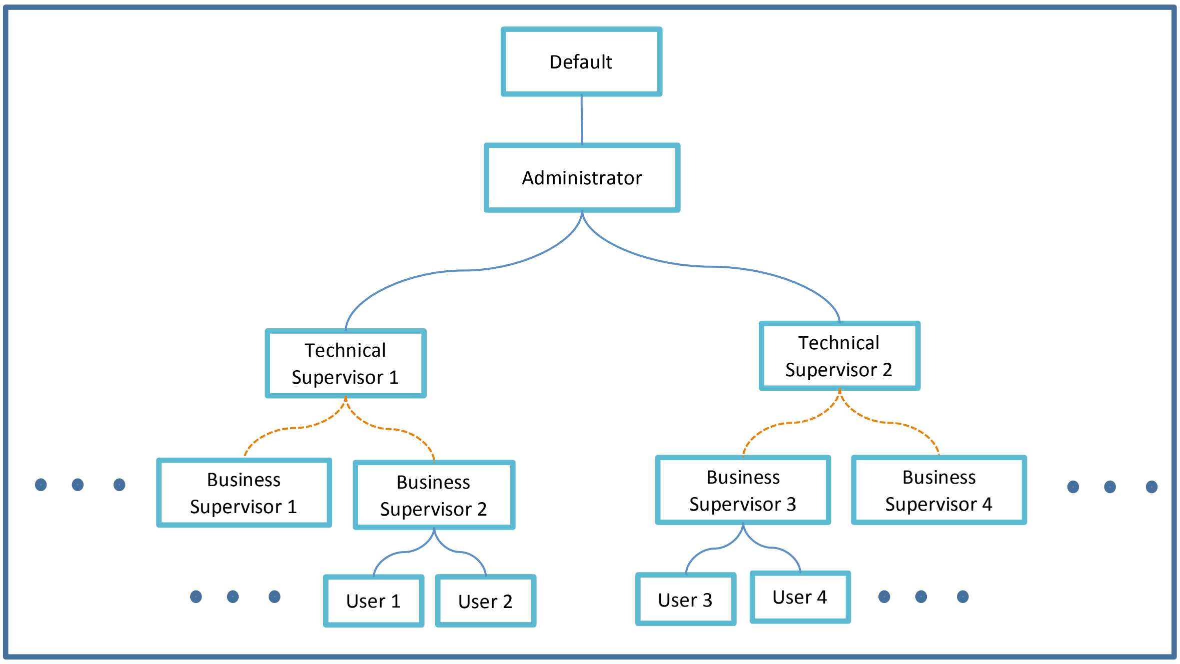

Figure 2. Proposed versioning schema

As illustrated in this figure, different versions have a hierarchical structure. User's versions are at the bottom of the hierarchy. When a user modifies an object, business supervisor can check them by comparing his/her version with the version of the related user. If the changes are valid from a business point of view then the changes will be transfered from user's version to the business supervisor version. Afterwards, technical supervisor can check the changes to assure that changes are applied correctly and GIS ready roles are observed. If this tow level of quality control passed, all the changes will be transfered upward to the administrator version. Finally, the versions of different users will be synchronized with transferring changes downward from administrator version to all other users.

4. IMPLEMENTATION

To illustrate the application of the model in a real world problem, we implement it in the Municipality of Shahinshahr city in the central part of Iran. The data were collected by the Municipality in digital format at the scale of 1:2000, totaling 20 sq. km. of area.
The framework for incremental and shared spatial data updating is developed as a client server GIS tool. Two packeges are developed for supervisors and editors. The first one developed as a stand alone system using VB.net and ESRI Engine Core. The second one developed using C++ as an extension of Autocad Map 3D software.
As illustrated in figure 3, the supervisors's system starts providing user with a wizard that enabling him/her to walk through some settings regarding connections to the server. Once the user sets his/her preferences with this wizard, the information is saved in his/her profile which will be used in the next usage of the system.

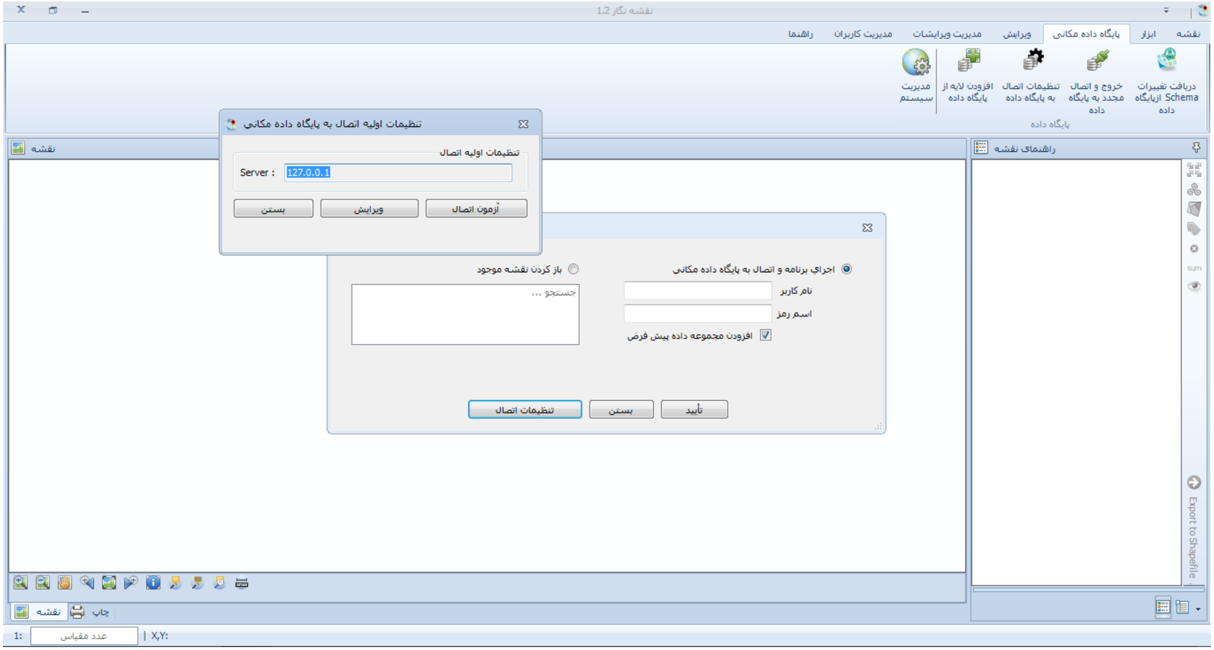

Figure 3. suppervisor's system

After providing user and password the system lists all the layers user have access where user can use version comparison tools to chekc the edits user have done. Figure 4 illustrates the panel for comparing versions.

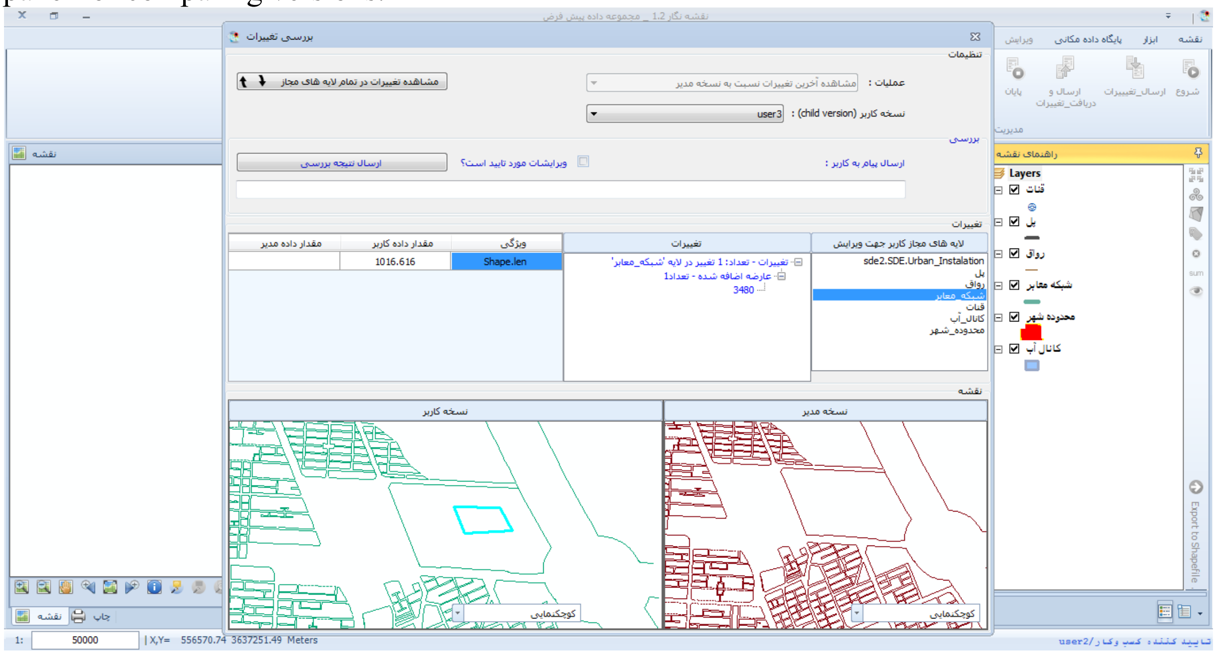

Figure 4. version comparision panel of supervisor's system

Figure 5 illustrates the extension developed over AutoCAD Map 3D. In this interface, after authentication passed user will have access to the related layers and can observe and/or edit them.

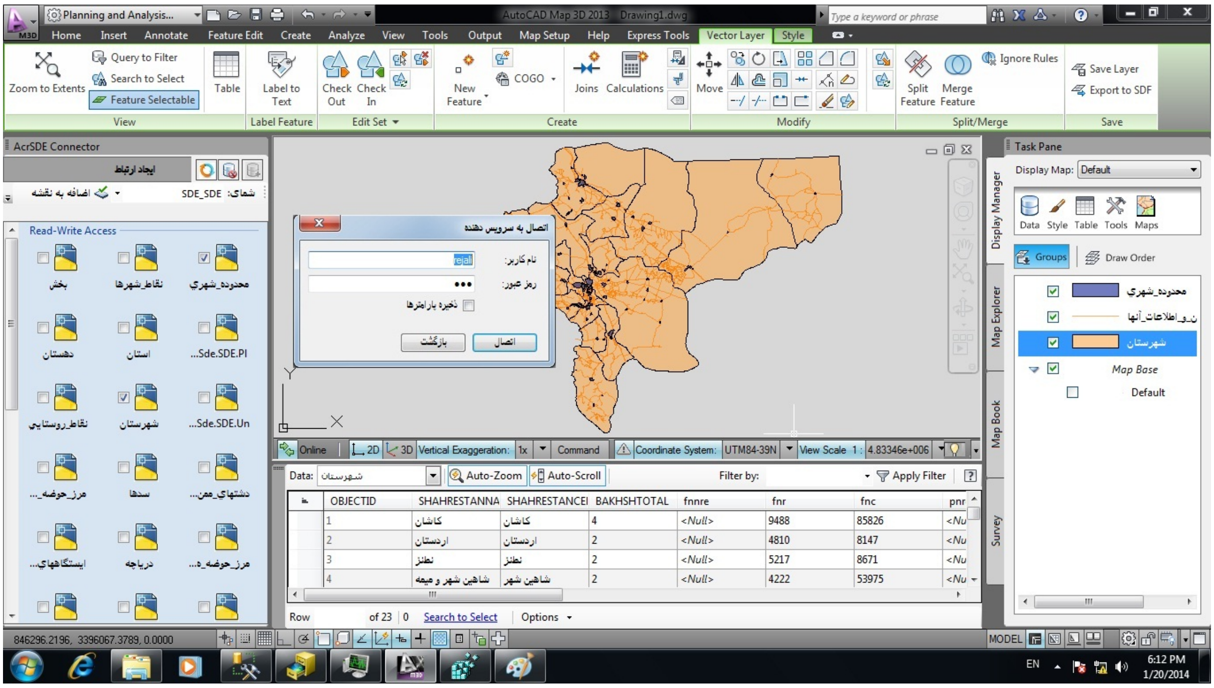

Figure 5. Spatial data sharing extension to AutoCAD Map 3D

5. CONCLUSION

This paper has proposed the theoretical basis for a model for incremental map updating in a shared environment. The model is based on distributed spatial database and used versioning as bases to make it possible for simultaneous updating. The model has a hierarchical structure with four levels. At the top, there is administrator which defines other users and synchronizes changes between different levels. In the second level there is technical supervisor which checks the changes to be GIS Ready. In the third business level supervisor check edits to assure that they are valid from a business point of view. And finally in forth level users have access to the data for editing, analysis or just observing and querying purposes.

The implemented model brings the capabilities of spatial data sharing within a distributed spatial database framework for incremental map updating which enhances the existing map updating systems in a shared environment. Our implementation is client server based and is designed for common users within organizations. Implementation of the system as a standalone system for supervisors and an extension for AutoCAD Map 3D simplify the usage of the system which is a critical success factor. The model implemented and tested in municipality of Shahinshahr city in Iran in which illustrates the efficiency and usefulness of the system. Future research is needed to investigate usability and usefulness of the system by conducting more tests addressing both the user interface design and the suggested method.

REFERENCES

Donaubauer , A. (2005). A multi-vendor Spatial Data Infrastructure for Local Government Based on OGC Web Services. *From Pharaohs to Geoinformatics.* Cairo, Egypt.

ESRI. (2003). *White Paper: Spatial Data standards and GIS Interoperability.* New York: ESRI Press.

Masser, I., Rajabifard, A., & Williamson, I. (2008). Spatially enabling governments through SDI implementation. *International Journal of Geographical Information Science*, 22 (1), 5-20.

Mohammadi, H., Rajabifard, A., & Williamson, I. (2009). Enabling Spatial Data Sharing through Multi-source Spatial Data Integration. *Proceedings of the GSDI 11 World Conference.* GSDI Association.

Onsrud, H., & Rushton, G. (1995). Sharing Geographic Information: An introduction. *Sharing Geographic Information* (pp. xiii-xviii). New Jersey: Center for Urban Policy Research.

BIOGRAPHICAL NOTES

Saeed Nadi is Assistant Professor and head of the Geomatics and Geospatial Information Systems (GIS) department at the Islamic Azad University, Maybod Branch. He received his Ph.D. in GIS from Tehran University. His research interests include the spatial database, spatial data sharing, temporal spatial data modeling, and Real-time routing.

CONTACTS

Dr. Saeed Nadi
Department of Surveying and Geomatics Engineering, Maybod Branch, Islamic Azad University
Maybod
Iran
Tel. (+98)-913-317-1781
Fax (+ 98)-311-7934514
Email: snadi@ut.ac.ir