Federated Data Model to Improve Accessibility of Distributed Cadastral Databases in Land Administration

Arbind Man TULADHAR and Mostafa RADWAN (The Netherlands) and Fatma Abdel KADER and Samir EL-RUBY (Egypt)

Key words: Cadastre, Land Administration, Information System, Federated Data Model, Spatial Data Infrastructure, GIS

SUMMARY

Geo-Information System manages a large volume of data concerning tenure security and market value and use. These data play dominant roles to the economical and environmental development of our society. Currently, such data are usually managed by many agencies in a distributed and heterogeneous processing environment in land administration within their organisational mandates. In such situation, the users such as individuals, notaries, real-estate brokers, surveyors, planners, municipalities and others face tremendous challenges in getting these datasets and integrating them into a single synchronous, consistent dataset using available GIS software.

This paper proposes the use of a federated data model (FDM), which acts as integrating layers on the top of the existing cadastral database systems distributed at the different locations. If such FDM is implemented in a single synchronous, consistent federated database system, the users would be relieved from the complex tasks of integration and yet get a consistent dataset for their uses at lowest possible cost.

For this purpose we firstly discuss aspects of geo-information management emphasizing on data management, users and sharing of cadastral data. Then we briefly review the concepts on federated database management system (FDMS) and its architecture. We then offer the federated data model approach for the purpose of land administration. In this content, "three-level federation architecture" is proposed. Then we briefly discuss the situation of land administration in Egypt. Finally the steps for federating distributed cadastral databases are presented. Preliminary find shows that we need to carry out such modeling practice specific to local situation to come up with an appropriate federated data model.

Federated Data Model to Improve Accessibility of Distributed Cadastral Databases in Land Administration

Arbind Man TULADHAR and Mostafa RADWAN (The Netherlands) and
Fatma Abdel KADER and Samir EL-RUBY (Egypt)

1. INTRODUCTION

One of the new emerging technologies in Geo-Information Management is today the concept of federation concept for improving access to the distributed databases in land administration and spatial data infrastructure. In this paper, we propose to develop such conceptual model in three-level federation architecture for integration of distributed and autonomous data using a mediation class in object oriented UML/OCL environment.

This paper firstly discusses aspects of geo-information management emphasizing on data management, users and sharing of cadastral data. Then we briefly review the concepts on federated database management system (FDMS) and its architecture. We then offer the federated data model approach for the purpose of land administration. In this content, "three-level federation architecture" is proposed. Then we briefly discuss the situation of land administration in Egypt. Finally the steps for federating distributed cadastral databases are presented. Preliminary find shows that we need to carry out such modeling practice specific to local situation to come up with an appropriate federated data model.

2. GEO-INFORMATION MANAGEMENT OF DISTRIBUTED CADASTRAL DATABASES

From the viewpoint of land administration, land has three main characteristics i.e. ownership, value and use. While it is considered as immovable physical object that are traded, the landowner may view it rather as a cultural heritage with social values in the form of emotional attachment to it. Real Estate experts consider it as economic object in which investments are put for its better use and hence enhancing its higher productivity for the better economy of a country.

In market driven economies, the management of the land ownership, value and use and their associated geo-information resources must operate within an integrated land policy framework (Dale and McLaughlin, 1999). For example, land markets are supported by a clear legal basis that is administered by regulatory authorities. Land and its associated buildings are traded according to their market value; they can be bought and sold, transferred from one owner to another or leased. The manner in which land or buildings may be used is, however, controlled by physical planning laws. Planning regulations affect the price that a purchaser will be willing to pay for any property since the permitted land use directly affects its market value (Tuladhar, 2005).

In many countries, the above activities are often treated in the fragmented manner for land management in the management of geo-information related to the ownership, economic value and use. The management of land ownership including cadastral surveying and mapping is often being controlled at a central government level, while the land use usually is at disposal of the municipal or local authority level. However some variations also occur in case of land value. Even in the case of land ownership registration and cadastre, more than one agency is usually involved.

2.1 Cadastral Data Management

In developing countries many cadastral data namely land ownership, cadastre, topography and property tax are organised and distributed under the different agencies including tax information. Access to these data sets is still very difficult due to the fact that they are in the analogue forms.

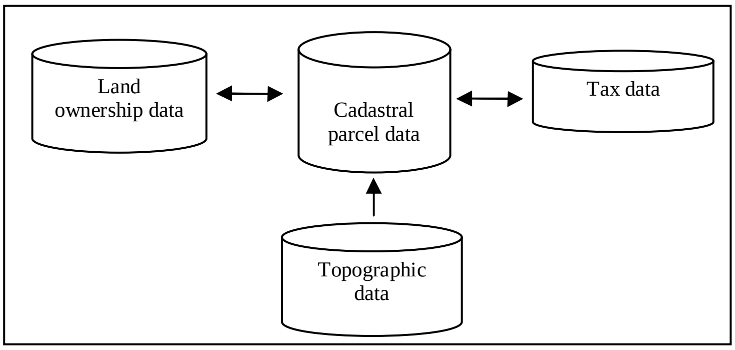

Figure no.1: Four Main Cadastral Data Sets in Land Administration

The figure no.1 shows four main distributed cadastral databases that are often found in land administration in many developing countries like Egypt. In a traditional cadastral system, information about property value and use are usually recorded either in cadastral parcels records or tax records. Since maintenance of these data sets is time-consuming and tedious in analogue forms, many of them normally lack reliability and quality of these data is hence questionable to the user community.

Many countries are now starting to convert the cadastral data into the structured digital databases for the quality products and services. Geo-ICT including web technology now allows us efficient and easy access to these data at the place where they are required. In this way, location does not matter where the data are stored. But from the management perspectives, the data are better kept at the location where the data is easily gathered and maintained.

2.2 Users and Sharing of Cadastral Data

Integrated cadastral data is required by the large variety of users such as municipalities, real estate brokers, provincial governments, insurance companies, financial agencies (including bank), utility companies who provides public services, various ministries, courts, individual citizen, notaries, etc. Here market orientation is considered as influential factor in the customer satisfaction, because in order to perform well, organisations like Egypt Survey Authority (ESA) needs relevant and timely information about the markets i.e. their customers and stakeholders or competitors. Because opportunities and threats continuously change e.g. due to the move made by stakeholders or competitors, the emergence of new technology, or shifts in customers' preferences and behaviours, the market must continuously be surveyed. In land registration and cadastral systems, we argued that the continuous stream of market data need to be collected, interpreted, distributed among organisation members, and be adequately utilised and exploited to stay competitive in the market (Tuladhar and Molen, 2003).

In order to satisfy the market needs, the data must be reliable and timely accessible to all users. In order to minimise data duplication, data sharing partnership between data producers are coordinated so that there are least conflicts on their data semantics. General quality aspects of the cadastral databases are related to completeness, actuality, accurateness, consistency (among the various databases) and accessibility. In understanding data sharing concept, the integration of data from various databases is a priority so that the users get consistent results with a least effort on data handling.

3. FEDERATED DATABASE MANAGEMENT SYSTEM (FDMS) AND ITS ARCHITECTURE

A federated database management system (FDBS) is a collection of cooperating but autonomous component database systems (DBSs). The component DBSs are integrated to various degree. A component DBS can participate in more than one federation, yet it continues its local operations. The integration of the component DBSs may be managed by either by the users of the federation or by the administration of the component DBSs. The amount of integration depends on the needs of federation users and desires of the administrators of the component DBSs to participate in the federation and share their databases.

3.1 Distribution, Heterogeneity and Autonomy

A FDBS consisting of multiple DBSs may be characterized along three orthogonal dimensions: distribution, heterogeneity and autonomy.

Distribution: Data may be distributed among multiple databases, stored on a single computer system or on multiple computer system. Benefits of data distribution, such as increased availability and reliability as well as improved access times, are well known (Sheth and

Larson, 1990). But in the case of FDBS, much of data distribution is due to the existence of multiple DBSs before an FDBS is built.

Heterogeneity: Many types of heterogeneity are due to technological difference, for example, differences in hardware, system software and communication system. In database environment, there are mainly two kinds: heterogeneities due to the differences in DBMSs and Semantic heterogeneity. The former is largely in differences in data models which have direct impacts in data structure, constraints and query languages.

Semantic heterogeneity occurs where there is a disagreement about the meaning, interpretation, or intended use of the same or related data. Semantic heterogeneity is broadly classified into four groups (Bishr, Molenaar and Radwan, 1996 and Bishr, 1997):

- differences in hierarchies
- differences in classes
- difference in geometry
- difference in attributes lists and domains

Autonomy: The organisational entities that manage different DBSs are often autonomous. That means that DBSs are often under separate and independent control. Those who control a database are often willing to let others share the data only if they retain control. In this respect, we can distinguish three types of autonomy. First type of autonomy is design autonomy that refers to the ability of a component DBS to choose its own design regardless of the data being managed, data models and the semantic interpretation of data which greatly contributes to the problem of heterogeneity. Secondly, execution autonomy refers to the ability of a component DBMS to execute local operations without interference from external operations. Third one is association autonomy which means that a component DBS has the ability to decide whether and how much to share its functionality and resources with others. This includes the ability to associate or disassociate itself from the federation and the ability of a component DBS to participate in one or more federation.

FDBS can be classified into loosely coupled and tightly coupled based on who manages the federation and how the components are integrated. An FDBS is loosely coupled if it is user's responsibility to create and maintain the federation and there is no control enforced by the federated system and its administrations. A federation is tightly coupled if the federation and its administration have responsibility for creating and maintaining the federation and actively control the access to component DBSs.

A federation is built by a selective and controlled integration of its components. The activity of developing an FDBS results in creating a federated schema upon which operations (i.e. query and update) are performed. Having a federated scheme (or federated data model) helps in maintaining uniformity in semantic interpretation of the integrated data.

3.2 Federated Database Architecture

The general architecture ensuring the different databases autonomy and preserving independent management including handling and administration is presented in the figure no. 2 adopted from Benchikha, Boufaida and Seinturier (2001).

Each database in the federation have its own local schema which are translated into common data model at the component level which is then exported to export schema representing part of databases relevant for federation. The complete integration is carried out at the federated level.

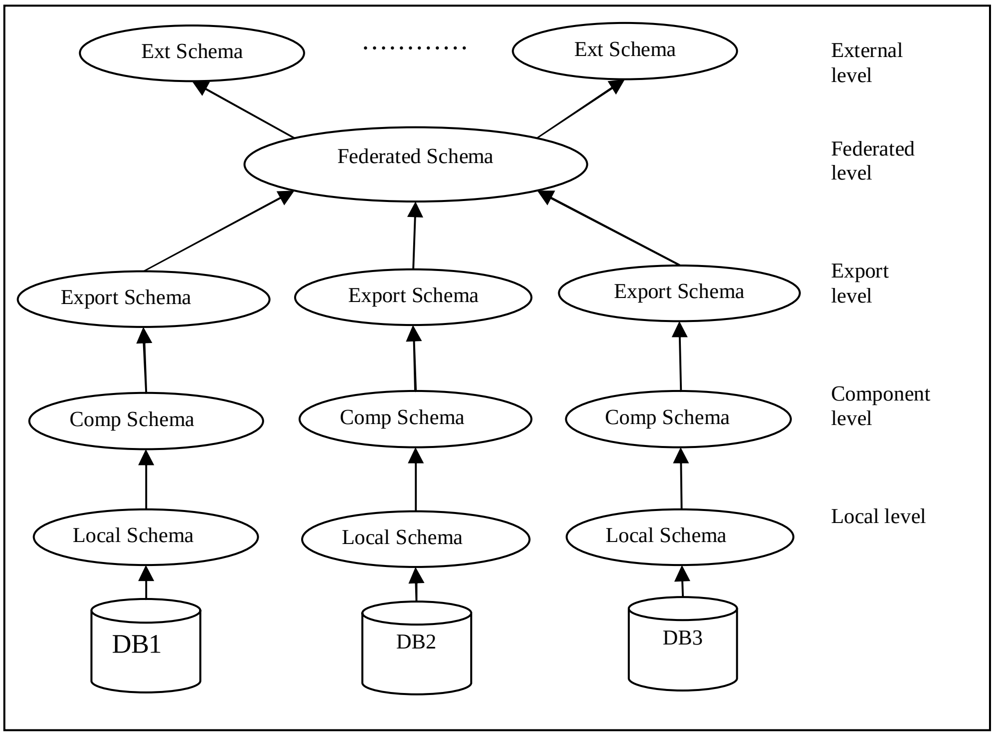

Figure no. 2: Federated database architecture (Benchikha, Boufaida and Seinturier, 2001)

Finally, the external level as shown in the figure no.2 provides relevant information for one of its users through external schemas. However, all levels are not necessarily required. But the critical issue here is to map export schemas to the federated schema (Heimbigner, 1985).

In the process of mapping between the export schema and federated schema, Bishr (1997) in his research work, introduced interesting steps resolving semantic heterogeneity by defining common ontology for context to share information. See the figure no. 3. This approach includes the definition of common ontology and mapping its elements on federated schema based on context information. Export schema corresponds with the common ontology and integrated to the federated schema.

It can be conclude that the approach is able to provide semantic resolution using the common ontology or context information without directly affecting the underlying database.

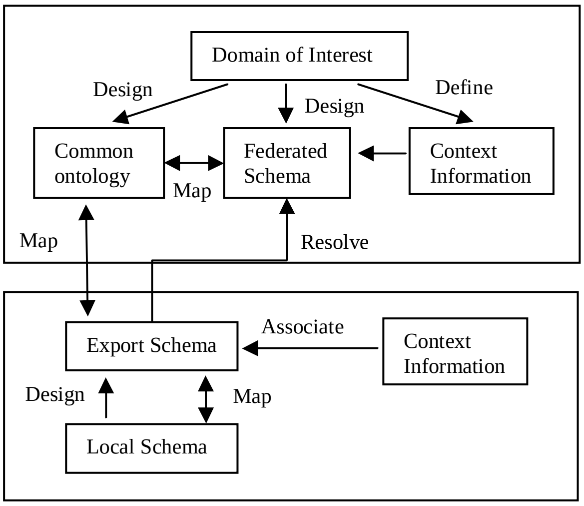

Figure no. 3: Mapping between Export schema and Federated schema (Bishr, 1997)

4. FDM APPROACH IN LAND ADMINISTRATION

In the field of Land administration, the cadastral databases are designed and implemented by the different organisations for a variety of purposes under the different legislation. Therefore, these databases use different schemas expressed by different data models to describe real-world objects. For example land ownership database is normally designed for the purpose of land transaction, while cadastral database is mainly to identify land parcel boundaries with area calculation supporting subdivision and taxation. Although land use and value are generally covered under these databases, they have to support local requirements under the constraints imposed by the local rules. As discussed in section 3.1, these databases may have a number of semantic heterogeneity.

The approach shown in figure no.3 uses common ontology and requires precise design for domain of interest with certain context. Alternatively, we propose here to use “three-level federation architecture” for land administration as shown in the figure no. 4. This architecture is based along the lines of traditional three-level architecture in database design, where user views are on the top of a conceptual schema of federation (federated schema) (Balsters, 2003b). The federated schema is derived using a mediator and a combination of internal or component schema inside a component frame which is in fact a collection of autonomous databases.

The realization of “three-level federation architecture” is also very much encouraged due to the availability of UML/OCL data model. UML is the *de facto* standard language for analysis and design of information systems in object oriented framework (Warmer and Kleppe, 2003). The notion of derived class facilitates as a means to integrate certain classes and constraints which need to be shared in a federated context. A mediator class is used for deriving federated schema from the component frame (Balsters, 2002).

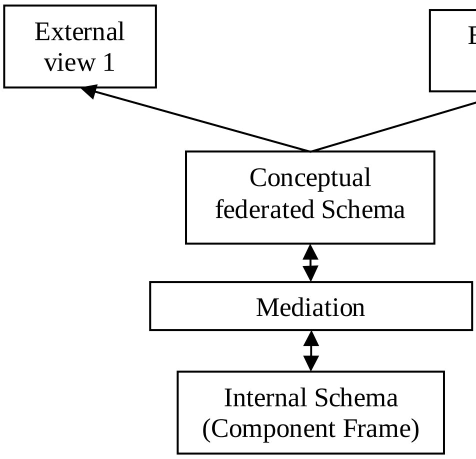

Figure no. 4: Three-level Federation Architecture based on Mediation

Integration of the internal schemas into one conceptual federated schema is usually done by systematically resolving the following conflicts (Balsters, 2003a):

- renaming (homonyms and synonyms)
- data conversion (different data types for related attributes)
- default values (adding default values for new attributes)
- missing attributes (adding new attributes in order to discriminate between certain class objects)
- subclassing (creating a common superclass and sequent accompanying subclasses)

Balsters (2003a and 2003b) argued that these conflicts can be resolved within a mediation class based on the principle of the tightly coupled approach to database integration and integrated database holds the union of the data in the source databases in the component frame.

5. LAND ADMINISTRATION IN EGYPT

In Egypt, the operations of land registration and cadastre are currently conducted by two agencies. The Ministry of Justice is responsible for legal transaction of private land ownership. Under this mandate, Real Estate Department (RED) of this ministry carries out daily transaction and maintains the land ownership record. Egyptian Survey Authority (ESA) under the Ministry of Public work and Irrigation is responsible for cadastral surveying and mapping component, and maintains geo-information about land parcel records including topographic data. Both agencies perform their tasks in accordance with the law 114/1946. The cadastral system is fiscal in nature and the ESA regularly supplies cadastral information to the Ministry of Finance for taxation purposes. The current procedure in Egypt for registration of land ownership data is not compulsory both for rural and urban areas. Cadastral surveying and mapping is rather complex due to the high accuracy requirements.

Since 1923, ESA is able to register and map 53% of agricultural land in rural areas. Currently, these databases are in analogue forms and the efforts to convert into the digital forms are taking place (El-Abedein, 2000 and Saleh, 1995)). Due to its limited cadastral coverage in the country in digital forms, the market economy in Egypt is severely affected. Increasing urbanisation, growing encroachment on the state owned land and the need to reclaim desert land have made completion of the cadastre a national priority including urban areas (Ibrahim and Hanigan, 1994).

In view of creating a new dynamic real estate market, the government has recently created the State Owned Land Office to keep the record of all government land and to create scenarios for investors over these lands (i.e. for tourist development, mining, agriculture, etc.).

In the current environment, land information is thus fragmented into the four main agencies making them to access difficult in an integrated fashion. But the components of these data would be required to share and integrate in delivering the data and services to the clients/users. The exchange and integration of these data sets among these agencies is currently done through analogue procedures.

5.1 Federating Distributed Cadastral Databases

As we indicate in section 5, there are land ownership, state owned land data, cadastral data, topographic data and tax data maintained by the real-estate department, state-owned land office, ESA and tax department respectively. There is good interaction between the ESA, RED and tax department at provincial and district levels for the processes of land transaction and cadastral surveying for agricultural land. These datasets are maintained and stored at their respective department at provincial levels.

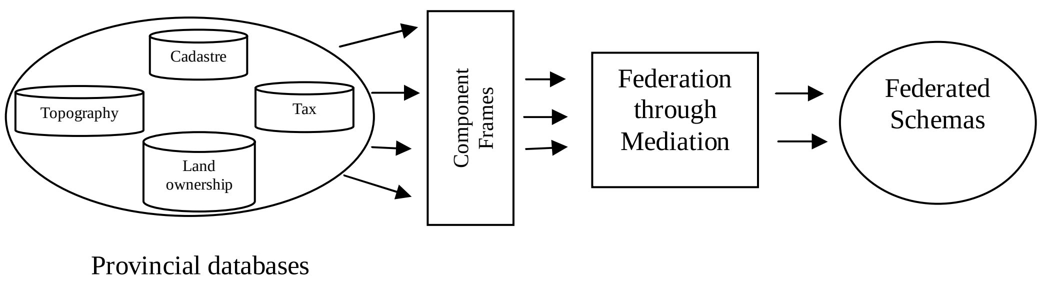

Figure no. 5: A proposed approach for federating provincial databases

The figure no.5 shows a proposed approach for federating provisional databases in Egypt using component frames and mediation class in UML/OCL environment. The following steps may be adopted for developing a federated schema or model.

- Component Frame is specified under UML model for all provincial databases
- Semantic heterogeneity is analyzed for detecting conflicts
- Introduce an explicit Class Mediator
- Construct an integrated schema
- Resolve conflicts via mediator class between component Frame and integrated schema (using suitable conversion functions).

5.2 Advantages of the Approach

There are several advantages in adopting the approach shown in the figure 6. Firstly the integration process does not affect individual databases at the provincial offices. These offices can continue working independently to satisfy the requests of their local users. The component frame means that local administration maintains the control over their systems and yet provides access to their data by the global users at the federation level. Such idea is also now emerging in spatial data infrastructure to facilitate reliable and timely access to spatial data

Since the architecture is of three-level, its implementation would be possible in Client/server environment, where a middle layer can take up federation task of the various component databases.

Preliminary findings show that the approach is promising. However, detail investigation on resolving conflicts via mediation in UML/OCL environment would be recommended specific to local situation and local constraints for developing appropriate federated schemas in land administration.

6. CONCLUSIONS

This paper discusses various approaches to access various databases using federated system which is one of new emerging technology for geo-data management. The paper proposes "three-level federation architecture" for land administration, and models federated system using UML/OCL. However, we need to carry out such modeling practice specific to local situation to come up an appropriate federated model. The issue of access control still needs to be investigated before implementation.

REFERENCES

Balsters, H., 2003a. A UML/OCL framework for design of mediated data federations; SOM Research Series 03A17, University of Groningen, 2000.

Balsters, H., 2002. Derived classes as a basis for views in UML/OCL data models; SOM Research Series 02A47, University of Groningen, 2002.

Balsters, H., 2003b. Object-oriented modeling and design of database federations; SOM Research Series 03A18, University of Groningen, 2003.

Benchika, F., Boufaida, M. and Seinturier, L., 2001. Integration of the viewpoint mechanism in federated databases. Proceedings of the 2001 ACM symposium on Applied computing, Las Vegas, Nevada, United States.

Bishr, Y.A., Molenaar, M. and Radwan, M.M., 1996. Spatial heterogeneity of federated GIS in a client/server architecture. Proceedings of Conference GIS World'96, Vancouver, Canada; on CD-ROM

Bishr, Y. 1997. Semantic aspects of Interoperable GIS. Phd thesis, ITC publication number 56, Enschede, the Netherlands.

El-abedein, H. Z., 2000. Toward a homogeneous cadastral map in Egypt: Map conversion process. Final Project, ITC, August, 2000.

Heimbigner, D. 1985. A federated Architecture for Information Management. ACM Transaction on Office Information Systems, Vol. 3, No. 3, July 1985.

Ibrahim, M. M. A. and Hanigan, F. L., 1994. Completing Egypt's Agricultural Cadastre. Egyptian Survey Authority, 1994.

Saleh, A. El-din, 1995. Temporal aspects in the framework of cadastral information system in Egypt. Unpublished M.Sc. thesis, ITC, June 1995.

Sheth, A. P. and Larson, J. A., 1990. Federated Database Systems for Managing Distributed, Heterogeneous and Autonomous Databases. ACM Computing Surveys, Vol. 22, No. 3, September 1990.

Tuladhar, A. M., 2004. Parcel-based Geo-Information system: concepts and guidelines. Phd thesis, ITC Dissertation series no. 115, Enschede, the Netherlands.

Tuladhar, A. M. and van der Molen, P. 2003. Customer satisfaction model and organizational strategies for land registration and cadastral systems. 2nd Cadastral congress, September 19-21 2003, Krakow, Poland.

Warmer, J. and Kleppe, A. 2003. The Object Constraint Language: Getting your models ready for MDA. Addison-Wesley

BIOGRAPHICAL NOTES

Dr. Arbind Man Tuladhar is assistant professor of Land Administration and Geo-Information Management at the department of Urban and Regional Planning and Geo-Information Management, International Institute for Geo-Information Science and Earth Observation (ITC), the Netherlands. His current research interests focus on land tenure, land administration, cadastral system, system modeling and development, and spatial data infrastructure.

Dr. Mostafa Radwan is associate professor of Geo-Information Infrastructure at the department of Urban and Regional planning and Geo-information Management, International Institute for Geo-Information Science and Earth Observation (ITC), the Netherlands. His current research interests focus on modeling, developing and managing multi-level Geo-Information Systems and Spatial Data Infrastructure. His email is radwan@itc.nl

Eng. Fatma Abdel Kader is director of cadastral survey and mapping at Egyptian Survey Authority, Cairo, Egypt. Her email is fatmaabdelkader@hotmail.com

Eng. Samir El-Ruby is project director for Egyptian Cadastral Information Management Project (ECIM) to develop digital cadastral system for the province of Behera at the Egyptian Survey Authority, Cairo, Egypt. His email is s_elrouby@hotmail.com.

CONTACT

Dr. Arbind Man Tuladhar
Assistant Professor
International Institute for Geo-Information Science and Earth Observation
P.O. Box 6, 7500AA
Enschede
THE NETHERLANDS
Tel. + 31 (0) 53 4874312
Fax + 31 (0) 53 4874575
Email: Tuladhar@itc.nl
Web site: http://www.itc.nl