

Chapter 39

Managing Schema Evolution in a Federated Spatial Database System

Xiaoying Wu

Curtin University, Australia

Geoff West

Curtin University, Australia

Jianhong (Cecilia) Xia

Curtin University, Australia

Lesley Arnold

Landgate, Australia

Bert Veenendaal

Curtin University, Australia

ABSTRACT

A Federated Spatial Database System (FSDBS) is the integration of multiple spatial data sources and the realisation of effective spatial data sharing. These are becoming increasingly popular as more and more spatial and non-spatial datasets are integrated, especially those across a number of independent organisations. However, in a FSDBS environment, database schemas are subject to change and the management of these changes is complex and inefficient. This is because schema changes in one local database will invalidate applications built against this local schema, but also applications built against the global schema.

The research is motivated by developments in the Shared Land Information Platform, built by Landgate in Western Australia as a Spatial Data Infrastructure enabler that has been running since 2007. The more than 350 datasets in SLIP are from and controlled by 20 organisations with queries built over the different datasets. Changes in the various databases require schema updating to be streamlined.

In this chapter, an Automatic Schema Evolution Framework is explored and developed to more effectively manage schema evolution in a FSDBS. This framework provides a Schema Element Dependency Meta-Model, a set of Schema Change Templates, and incorporates view generation, view rewriting, and query rewriting as solutions. These developed methods ensure applications can accommodate schema changes and hence remain valid.

DOI: 10.4018/978-1-4666-2038-4.ch039

INTRODUCTION

The traditional database can be described by schemas that define relations, attributes, etc. A spatial database schema is an extension of the traditional schema to include spatial descriptions and behaviours required for a spatial database (Yeung & Hall, 2007). Spatial database schemas, like traditional schemas, are subject to change or evolution due to changes in perception of reality and application requirements. Changes can also occur as a consequence of integration with other systems, compliance to new regulations and the implementation of new security requirements.

The task of managing schema evolution is to ensure schemas, data and application consistency and integrity after change. Schema changes often result in the applications built against the schemas being invalid.

In a Federated Spatial Database System (FSDBS), in which spatial means geographically spatial, the spatial data are shared by multiple organizations and applications across a number of databases. A local database schema change will affect not only the applications locally but also applications built on the federated schemas. Effective management is problematic and there are several challenges to managing schema evolution in a FSDBS including:

- How to evaluate the impact of schema changes on local and global databases;
- How to propagate the schema changes across databases in a federated environment; and,
- How to automatically rewrite run time queries, and rewrite and generate views (spatial views and traditional views) so that the applications are immune to schema changes.

The objective is to explore and develop an Automatic Schema Evolution (ASE) framework for managing schema evolution in a FSDBS. The main contribution of this chapter is the proposal

of semi-automatic methods that, in future work, will lead to the development of fully automatic methods.

Of particular interest is the Shared Land Information Platform (SLIP) built by Landgate (authoritative agency for land information in Western Australia) as a Spatial Data Infrastructure (SDI) enabler (2011). SLIP went into production in 2007 and consists of more than 350 datasets from 20 different agencies. As such it is a federated database and queries are built over the different datasets. Schema updating is needed because the datasets are agency owned and controlled allowing local changes to occur.

For schema evolution, the mismatch between evolved schemas and queries need to be overcome. Queries here include queries generated by applications and stored queries (views and spatial views) as seen by users. The ASE framework defines the components and methods for automatic run time query and view rewriting, as well as new view generation when schemas change.

BACKGROUND

Before explaining the proposed ASE framework, it is necessary to define various aspects required for the ASE Framework, explore the various issues and concepts involved in the framework, and critique the current state of the art in research into federated and spatial databases.

Spatial Database Management Systems

In order to model the spatial aspects of real world objects, a number of different approaches have been proposed and developed. For example, the MADS (Modeling Application Data with Spatio-Temporal features) model defines constructs and language to describe an application schema from different aspects including thematic and spatial-temporal data structures, and multi-representation aspects

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/managing-schema-evolution-federated-spatial/70467

Related Content

Modeling Hydrological Functioning of a Drainage Basin With Relation to Land Use Change in the Context of Climate Change: Ourika Watershed Case Study

Reda Rihane, Abdellatif Khattabi, Nabil Rifai and Said Lahssini (2019). *Geospatial Technologies for Effective Land Governance* (pp. 194-213).

www.irma-international.org/chapter/modeling-hydrological-functioning-of-a-drainage-basin-with-relation-to-land-use-change-in-the-context-of-climate-change/214488

IJAGR at the International Medical Symposium Geography 2022 Edinburgh, Scotland

Donald P. Albert (2023). *International Journal of Applied Geospatial Research* (pp. 1-5).

www.irma-international.org/article/ijagr-at-the-international-medical-symposium-geography-2022-edinburgh-scotland/318138

Biometric Authentication in Broadband Networks for Location-Based Services

George Kakaletis, Dimitris Varoutas, Dimitris Katsianis and Thomas Sphicopoulos (2009). *Handbook of Research on Geoinformatics* (pp. 422-432).

www.irma-international.org/chapter/biometric-authentication-broadband-networks-location/20426

A Comparison of Principal Component-Based and Multivariate Regression of Cardiac Disease

Fox Underwood and Stefania Bertazzon (2013). *Geographic Information Analysis for Sustainable Development and Economic Planning: New Technologies* (pp. 31-48).

www.irma-international.org/chapter/comparison-principal-component-based-multivariate/69047

Towards the Integration of Trajectory Information Sources for Semantic Conflicts Detection Purpose: A Trajectory Ontology Based Approach

Marwa Manaa and Akaichi Jalel (2017). *Handbook of Research on Geographic Information Systems Applications and Advancements* (pp. 488-519).

www.irma-international.org/chapter/towards-the-integration-of-trajectory-information-sources-for-semantic-conflicts-detection-purpose/170002