Chapter 3.20
Introducing the Elasticity of Spatial Data

David A. Gadish
California State University Los Angeles, USA

ABSTRACT

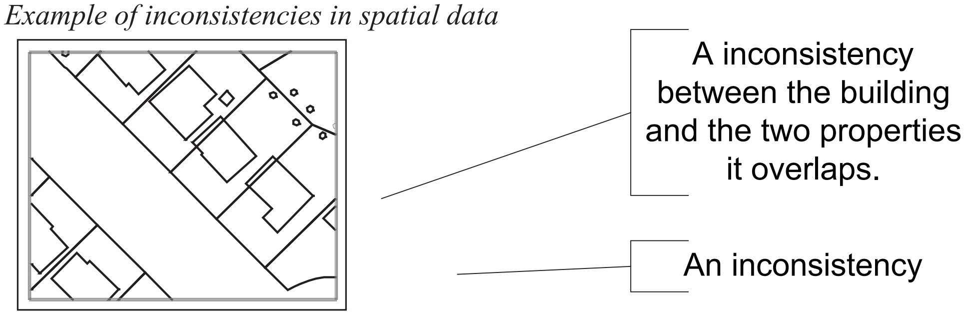
The data quality of a vector spatial data can be assessed using the data contained within one or more data warehouses. Spatial consistency includes topological consistency, or the conformance to topological rules (Hadzilacos & Tryfona, 1992, Rodríguez, 2005). Detection of inconsistencies in vector spatial data is an important step for improvement of spatial data quality (Redman, 1992; Veregin, 1991). An approach for detecting topo-semantic inconsistencies in vector spatial data is presented. Inconsistencies between pairs of neighboring vector spatial objects are detected by comparing relations between spatial objects to rules (Klein, 2007). A property of spatial objects, called elasticity, has been defined to measure the contribution of each of the objects to inconsistent behavior. Grouping of multiple objects, which are inconsistent with one another, based on their elasticity is proposed. The ability to detect groups of neighboring objects that are inconsistent with one another can later serve as the basis of an effort to increase the quality of spatial data sets stored in data warehouses, as well as increase the quality of results of data-mining processes

INTRODUCTION

Geographic Information Systems

Spatial databases, and most notably Geographic Information Systems (GIS) databases and data warehouses, have gained popularity in various business sectors in recent years. GIS or spatial databases may consist of vector, raster, and nonmapping business data (Abler, 1987). This article is focused on the vector data stored in these databases. GIS databases are comprised of

Figure 1. Example of inconsistencies in spatial data



vector-based data objects such as points, lines, and polygons (Aronoff, 1995). These elements are related to one another in terms of their location as well as in terms of their semantic interpretation (Clementini, Felice, & Van Ooserom, 1993; Egenhoffer, 1991).

Inconsistency of Spatial Data

Consistency is a measure of data quality of these databases and data warehouses. Inconsistency is the opposite of consistency (Kainz, 1991). Inconsistency can be measured in terms of the number of objects that relate to one or multiple objects that do not satisfy one or more rules that describe the data (Cockcroft, 2004; Lutz & Kolas, 2007; Xu, 2007). For example, it would be inconsistent to map a property boundary to the center of an adjacent road. Consistency is assessed using information contained within one or more databases. Spatial consistency includes topological consistency, or the conformance to topological rules (Kainz, 1991). An example of spatial inconsistency is illustrated in Figure 1, where two buildings slightly overlap their properties.

Detection of Inconsistencies in Spatial Data Warehouses

Detection of inconsistencies in spatial data is a key activity in an effort to manage the consistency of spatial data (Servigne, Ubeda, Zuricelli, & Laurini, 2000). It is useful for evaluating, maintaining, and enhancing spatial data quality in data warehouses. Since data in spatial databases is constantly evolving, there is an ongoing need for detection and adjustment of inconsistencies. This need is amplified as two or more spatial data sources are combined in a warehouse to increase the scope of data mining. It is also useful for enhancing the quality of results produced by data-mining operations.

Figure 2. The process of creating elasticity-based groupings