Chapter 76
Knowledge Extraction from Geographical Databases for Land Use Data Production

Hana Alouaoui
LTSIRS Laboratory, University of Tunis El Manar, Tunisia

Sami Yassine Turki
LTSIRS Laboratory, University of Tunis El Manar, Tunisia

Sami Faiz
LTSIRS Laboratory, University of Tunis El Manar, Tunisia

ABSTRACT

Our study focuses on the task of land use evolution in urban environment which is fundamental in revealing the territorial planning. It refers crucially to the use of spatial data mining tools due to their high potential in handling with spatial data characteristics. The results of our knowledge discovery process are spatial and spatiotemporal association rules referring to the land use and its evolution. Three proposals based on different knowledge extraction techniques are detailed. The first approach aims to extract spatiotemporal association rules by introducing time into the attributes. The second approach forecasts the extracted rules at different dates. The third approach is devoted to the mining of spatiotemporal association rules. This proposal looks for rules that relate properties of reference objects with properties of other spatial relevant objects. The extracted patterns are relationships involving the spatial objects during time periods. To prove the applicability of each approach, experimentations are conducted on real world data. The obtained results are promising.

INTRODUCTION

The land use is the manner in which the territory is covered and organized and whose human being has taken possession of the geographical area and fashioned it in order to live. This is distinct from the soil use definition referred to how land is used and exploited. But the two concepts have a major correlation

DOI: 10.4018/978-1-5225-7033-2.ch076

(Salem, 1996). The progress in the production of land use data has offered multiple approaches based on ground data collecting (topographical surveys, GPS, etc.) and remote sensing (aerial photography, satellite images, etc.). Therefore, various techniques have been proposed such as neural networks, cellular automata, Geographical Information Systems (GIS), entropy measurement, multiple regression, etc. The land use knowledge is a prerequisite for any planning or management action. Thereby providing a mapping of land use constantly updated becomes a serious issue. From a methodological point of view, several approaches can be useful in land data production. Two different axes are distinguished; the identification of a current land use type and the estimation of the land use projection.

If we put aside the ground approaches (surveying or GPS) that cannot be used on a large scale, the use of remote sensing is the most convenient for such purposes. Indeed, the use of satellite imagery especially in urban areas has been the subject of several studies (Devriendt et al., 2004; Nardinocchi et al., 2001). Lots of methods of knowledge production from satellite images such as automatic classification or supervised classification (Cornet et al., 2004; Renard et al., 1997), extraction of numerical models of surfaces (Binard et al., 2002), allow obtaining a knowledge on land use, specifically the occupation mode: empty, built, vegetable, mineral, roads, buildings, etc. Various studies have been designed to adapt specific methodologies reducing the effect of "salt and pepper" of pixels in the high resolution images such as; filtering and segmentation at different scales (Van de Voorde et al., 2004). Other methods propose to combine the high resolution of images with contextual analysis (Teeffelen et al., 2001) or even the use of theories issued from animal biology to contribute to the fusion or classification processes (Khedam, 2008). Additionally, a big attention has been devoted to the recognition and understanding of urban space. Such objective can be apprehended through the urban forms and their characteristics. Knowledge of such elements can provide insights on the characteristics of the ground dynamism (Isik et al., 2013; Weber et al., 1997). Zhang and Zhu (2011) as other authors point out the limitations of visual interpretation and automatic classification of images to produce information. They propose to use as knowledge the rules extracted from spectral and textural data. However, beyond these limits, some data on the built spaces function cannot be produced, e.g. we distinguish an industrial building from a house but we cannot say if a group of buildings is used for residential purpose or for offices.

The estimation of the land use projection is revealed as another emergent research axis. Hence, the exploration of spatiotemporal data has come to the fore with the objective to study the urban evolution (Boulila, 2012; Essid et al., 2012; Yang, 2006; Zeitouni, 2006). Spatiotemporal data mining applies new tools in analyzing very large spatiotemporal databases. However, it presents a number of challenges due to the complexity of spatial and temporal characteristics (Gidofalvi & Pedersen, 2006; Jouini, 2006; Kuntala & Raghavan, 2006; Otoo & Rotem, 2006). Additionally, applications dealing with spatiotemporal data are specifically addressed to cell phone networks, traffic management, sensor networks (Chen et al., 2006; Vuran et al., 2004) or tracking objects moving throughout regions (Mamoulis et al., 2004). Other approaches are devoted to study the identity based changes (Hallot, 2012; Hornsby & Egenhofer, 2000) as particular types of spatiotemporal data. The main idea in (Pouliot et al., 2014) is to detect cover changes from satellite-based earth observation. The detected changes are used to update maps. In (Neeti and Eastman, 2014), the concept of concatenation of multi-temporal images is introduced. Three different sets of precipitation are used as a case study. The approach presented in (Lyons et al., 2013), analyses the seagrass growth. Compared to traditional seagrass mapping and monitoring approaches, this method proves clear benefits. In (Dinga et al., 2014), a concept of mean length variability is proposed to compare the difference in spatial heterogeneity. Temporal changes in spatial heterogeneity are observed and they are a result of changes in the fraction of vegetation cover.