IDEA GROUP PUBLISHING

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA

Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

ITB10789

Chapter XII

Mining in Spatio-Temporal Databases

Junmei Wang, National University of Singapore, Singapore

Wynne Hsu, National University of Singapore, Singapore

Mong Li Lee, National University of Singapore, Singapore

Abstract

Recent interest in spatio-temporal applications has been fueled by the need to discover and predict complex patterns that occur when we observe the behavior of objects in the three-dimensional space of time and spatial coordinates. Although the complex and intrinsic relationships among the spatio-temporal data limit the usefulness of conventional data mining techniques to discover the patterns in the spatio-temporal databases, they also lead to opportunities for mining new classes of patterns in spatio-temporal databases. This chapter provides a survey of the work done for mining patterns in spatial databases and temporal databases, and the preliminary work for mining patterns in spatio-temporal databases. We highlight the unique challenges of mining interesting patterns in spatio-temporal databases. We also describe two special types of spatio-temporal patterns: location-sensitive sequence patterns and geographical features for location-based service patterns.

Introduction

The globalization of many industries and advances in wireless communication and global positioning systems have led to the development of spatio-temporal applications, such as applications dealing with moving objects, involving objects located in the space (for example, land parcels, whose characteristics may change in time) and dealing with objects which integrate the above two behaviors (for example, pollution phenomenon in the environment system). A spatio-temporal database embodies spatial, temporal and spatio-temporal concepts, and captures simultaneously the spatial and temporal aspects of data. The complex interactions among the objects are captured in the form of the past, present and future states in the modeled environment. With the widespread use of spatio-temporal databases, there is the increasing need for the discovery of interesting and previously unknown, but potentially useful, patterns in spatio-temporal databases.

Currently, spatial data mining and temporal data mining form two separate streams of research. Efforts are either focused on discovering space-sensitive patterns in the form of *spatial patterns* (Chawla, Shekhar, &Wu, 2000; Ester, Frommelt, Kriegel, & Sander, 2000; Han, Koperski, & Stefanovic, 1997; Shekhar & Huang, 2001; Wang, Yang, & Muntz, 1999); or time varying patterns in the form of *sequence patterns* (Agrawal & Srikant, 1996; Pei, Han, & Asl, 2001; Zaki, 1998; Yang, Wang, Yu, & Han, 2002; Mannila, Toivonen, & Verkamo, 1995; Garofalakis, Rastogi, & Shim, 1999). However, spatio-temporal databases contain the complex relationships that cannot be discovered by simply considering the temporal information or spatial information independently.

For example, it is not sufficient to know that "*sales are typically up in the months of February and December*," or that "*sales are high in Asia and North America*." Instead, it is more important to understand the trends in the context of spatial locations, such as "*sales in the Asia region peak in the month of February while sales in the North America continent peak in the month of December*." Further, the pattern "*motorists go to gas stations upon exit from a freeway with 80% likelihood*" can only be discovered when the mining algorithm takes into account the multi-states (that is, the past, present and future states) of the objects.

Data mining in spatio-temporal databases must consider the multi-states of the spatio-temporal data. It must integrate the spatial information and temporal information together to find meaningful spatio-temporal patterns. The knowledge of these spatio-temporal patterns allows one to develop more localized or customized business analysis and strategies, and have potential benefits for many applications, such as the Geographic Information System (GIS), environ-