



Chapter XI

Spatio-Temporal Prediction Using Data Mining Tools

Margaret H. Dunham, Southern Methodist University, Texas, USA

Nathaniel Ayewah, Southern Methodist University, Texas, USA

Zhigang Li, Southern Methodist University, Texas

Kathryn Bean, University of Texas at Dallas, USA, USA

Jie Huang, University of Texas Southwestern Medical Center, USA

Abstract

The spatio-temporal prediction problem requires that one or more future values be predicted for time series input data obtained from sensors at multiple physical locations. Examples of this type of problem include weather prediction, flood prediction, network traffic flow, and so forth. In this chapter we provide an overview of this problem, highlighting the principles and issues that come to play in spatio-temporal prediction problems. We describe some recent work in the area of flood prediction to illustrate the use of sophisticated data mining techniques that have been

examined as possible solutions. We argue the need for further data mining research to attack this difficult problem. This chapter is directed toward professionals and researchers who may wish to engage in spatio-temporal prediction.

Introduction

Forecasting future values for systems that contain both spatial and temporal features (spatio-temporal) is extremely complex. As an example, consider the problem of predicting precipitation at one location. The amount of previous rainfall in areas close to the target certainly affects this forecast. However, there are many other factors (temperature, time of day, wind direction, wind speed, and so forth) that impact the rainfall prediction. The area of spatio-temporal prediction has been the focus of much research in recent years (Deutsch, & Ramos, 1986; Dougherty, Corne, & Openshaw, 1997; Jothityangkoon, Sivapalan, & Viney, 2000; Kelly, Clapp, & Rodriguez, 1998; Pokrajac, & Obradovic, 2001; Roddick, Hornsby, & Spiliopoulou, 2000; Singh, Chaplain, & McLachlan, 1999). Due to the extreme complexity of predicting these future values, common practice is to utilize domain experts with extensive experience in both forecasting and the problem domain itself. For example, for flood prediction, the *National Oceanic and Atmospheric Administration (NOAA)* actually employs specialists whose job is to understand the history and specifics of predicting floods on one river. A different domain expert may be hired for a different river. Due to the extensive use of domain experts, spatio-temporal prediction is extremely expensive, and due to the complexity of the nature of the problems, prediction accuracy is often low.

In this chapter we argue for more data mining research into the development of sophisticated modeling, machine learning and prediction tools that can assist domain experts in solving spatio-temporal prediction problems. The interesting challenge is to see if these models can use all the data available without the need of expert intervention. This could be useful, since experts in the particular domain of interest – flood prediction, networks – often may not be experts in these data mining tools, which are drawn from statistics, genetic modeling, algorithmic heuristics and much more. But creating models general enough to be understood by non-experts has proven to be a difficult balancing act, especially when these complex models use many parameters and when there are many types of data to consider.

In this chapter we first introduce the problem and some previous research and solutions. Finally we summarize and make recommendations for future work.

19 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/spatio-temporal-prediction-using-data/29667

Related Content

Collaboration Matrix Factorization on Rate and Review for Recommendation
Zhicheng Wu, Huafeng Liu, Yanyan Xu and Liping Jing (2019). *Journal of Database Management* (pp. 27-43).

www.irma-international.org/article/collaboration-matrix-factorization-on-rate-and-review-for-recommendation/232720

A Study of a Generic Schema for Management of Multidatabase Systems
Shirley A. Becker, Rick Gibson and Nancy L. Leist (1996). *Journal of Database Management* (pp. 14-20).

www.irma-international.org/article/study-generic-schema-management-multidatabase/51169

Management of Large Moving Objects Databases: Indexing, Benchmarking and Uncertainty in Movement Representation

Talel Abdessalem, Cédric du Mouza, José Moreira and Philippe Rigaux (2005). *Spatial Databases: Technologies, Techniques and Trends* (pp. 225-250).

www.irma-international.org/chapter/management-large-moving-objects-databases/29666

An Overview of Fuzzy Approaches to Flexible Database Querying

Slawomir Zadrozny, Guy de Tré, Rita de Caluwe and Janusz Kacprzyk (2009). *Database Technologies: Concepts, Methodologies, Tools, and Applications* (pp. 135-156).

www.irma-international.org/chapter/overview-fuzzy-approaches-flexible-database/7906

Representing Classes of Things and Properties in General in Conceptual Modelling: An Empirical Evaluation

Graeme Shanks, Daniel Moody, Jasmina Nuredini, Daniel Tobin and Ron Weber (2010). *Journal of Database Management* (pp. 1-25).

www.irma-international.org/article/representing-classes-things-properties-general/42083