



## **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](http://www.igi-global.com/chapter/spatio-temporal-prediction-using-data/29667)

## Related Content

---

### ERP Systems Impact on Organizations

Jonas Hedman and Andreas Borell (2003). *ERP & Data Warehousing in Organizations: Issues and Challenges* (pp. 1-21).

[www.irma-international.org/chapter/erp-systems-impact-organizations/18552](http://www.irma-international.org/chapter/erp-systems-impact-organizations/18552)

### Design and Representation of Multidimensional Models with UML and XML Technologies

Juan Trujillo, Sergio Lujan-Mora and Il-Yeol Song (2005). *Advanced Topics in Database Research, Volume 4* (pp. 50-95).

[www.irma-international.org/chapter/design-representation-multidimensional-models-uml/4368](http://www.irma-international.org/chapter/design-representation-multidimensional-models-uml/4368)

### On the Distributed Database Union Operation

Rajiv M. Dewan (1992). *Journal of Database Administration* (pp. 24-35).

[www.irma-international.org/article/distributed-database-union-operation/51103](http://www.irma-international.org/article/distributed-database-union-operation/51103)

### INDUSTRY AND PRACTICE: An Empirical Investigation of the Effectiveness of Object-Oriented Database Design

Chetan Sankar and Debasis Rath (1994). *Journal of Database Management* (pp. 39-40).

[www.irma-international.org/article/industry-practice-empirical-investigation-effectiveness/51141](http://www.irma-international.org/article/industry-practice-empirical-investigation-effectiveness/51141)

### Web Data Warehousing Convergence: From Schematic to Systematic

D. Xuan Le, J. Wenny Rahayu and David Tanir (2009). *Database Technologies: Concepts, Methodologies, Tools, and Applications* (pp. 678-701).

[www.irma-international.org/chapter/web-data-warehousing-convergence/7938](http://www.irma-international.org/chapter/web-data-warehousing-convergence/7938)