Applied Time Series Analysis and Innovative Computing

Lecture Notes in Electrical Engineering
Volume 59

For other titles published in this series, go to
www.springer.com/series/7818

Sio-Iong Ao

Applied Time Series Analysis and Innovative Computing

Springer

Sio-Iong Ao
International Association of Engineers
Unit 1, 1/F, 37–39 Hung To Road
Hong Kong
publication@iaeng.org

ISBN 978-90-481-8767-6 e-ISBN 978-90-481-8768-3
DOI 10.1007/978-90-481-8768-3
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2010924436

To My Lovely Wife Choi, Wai-Ming

Preface

There are many reasons to analyze the time series data, for example, to understand the underlying generating mechanism better, to achieve optimal control of the system, or to obtain better forecasting of future values. Applied time series analysis consists of empirical models for analyzing time series in order to extract meaningful statistics and other properties of the time series data. With the advances in computer technology, nowadays huge amounts of time series data are stored in data warehouses. Different innovative computing techniques are needed to extract information from these datasets. Innovative computing paradigms can extract the patterns from the time series and to present it in such a way that can better our understanding of the structure, relation, and function of the subjects. The purpose of this book is to illustrate how to apply innovative computing paradigms for the applied time series analysis, with frontier application studies of the time series problems based on the recent works at the Oxford University Computing Laboratory, University of Oxford, the University of Hong Kong, and the Chinese University of Hong Kong. The monograph was drafted when the author was a postdoctoral fellow in Harvard School of Engineering and Applied Sciences, Harvard University. It provides a systematic introduction to the use of innovative computing paradigms as an investigative tool for applications in time series analysis.

In this book, innovative computing algorithms have been illustrated for solving some frontier problems in time series analysis. The book is organized as follows. In Chap. 1, it is the brief introduction to the applied time series analysis and the advances in innovative computing paradigms. The three real-word applications of innovative computing paradigms for time series problems and the contributions of these algorithms to the time series analysis are described briefly. In Chap. 2, we describe about the applied time series analysis generally. Time series analysis models including time domain models and frequency domain models are covered. In Chap. 3, we describe about the recent advances in innovative computing paradigms. Topics like computing algorithms and databases, integration of hardware, systems and networks, Internet and grid computing, and visualization, design and communication, will be covered. The advances of innovative computing for time series problems are also discussed, and an example of building of an innovative computing algorithm for some simulated time series is illustrated. In Chap. 4, we present the real-word application of innovative computing paradigms for time series problems.

The interdisciplinary innovative computing techniques are applied to understand, model and design systems for business forecasting. In Chap. 5, the second real-word application is for the analysis of the biological time series. Recurrent Elman neural networks and support vector machines have been outlined for temporal modeling of microarray continuous time series data sets. In Chap. 6, we present the last real-word application for the astronomical time series. Some innovative computing algorithms are described about how to classify the light curves of the quasars against light curves of the other stars.

In business organizations, forecasting is one of the most important activities that form the basis for strategic and operational decision. Traditionally, business time series forecasting has been dominated by linear methods. However, the linear models have serious limitation with problems of nonlinear relationships. It may be unsatisfactory to approximate the linear models for these nonlinear relationships. The importance and complexity of the business time series forecasting problem paves way for the importance of innovative computing paradigms. In Chap. 4, the interdisciplinary innovative computing techniques are applied to understand, model and design systems for business forecasting. Two business case studies are described to demonstrate the advantages of our proposed system. In global business, the interactions between different markets cause collective lead-lag behavior having special statistical properties which reflect the underlying dynamics. The internal structure of a complex system can manifest itself with correlations among its components. An innovative computing system of combining the vector autoregression and genetic algorithm with neural network is outlined how to take advantage of the lead-lag dynamics, to make the neural network forecasting process more transparent and to improve the prediction capability of the neural network.

The temporal patterns exhibited by biological time series are often complex, nonlinear or chaotic, nevertheless, the analysis are emerging as important means for many biological studies. To address the nonlinear properties of the biological time series, innovative computing algorithms are becoming important tools to study these systems. In Chap. 5, recurrent Elman neural networks and support vector machines are outlined for temporal modeling of microarray continuous time series data sets. In a single microarray experiment, the expression levels of as many as thousands of genes can be measured simultaneously. Thus, it can enable the genome-wide measurement of gene expression, and the construction of genetic network from gene expression time series with innovative computing approaches. An ensemble of the ENN and SVM models are described to better improve the transparency and robustness of the recurrent neural networks, and to further improve the prediction accuracy of the individual models. In order to provide the neural networks with explanation capabilities, a pedagogical rule extraction technique is considered for inferring the output of our proposed ensemble system. The gene regulatory network can be reconstructed satisfactorily with this hybrid innovative computing methodology. The proposed rule extraction technique provides the network inference of the ensemble. The reconstructed gene regulatory network enables us to better understand the dynamics of the underlying biological processes.

With the advances of the technologies for the sky surveys, massive amount of survey data become available. It is very helpful for the automatic and semi-automatic innovative computing methods in the classifications and detections of the astrophysical objects. In Chap. 6, the first section describes about the existing automatic and semi-automatic innovative computing methods for the comprehensive search of quasars. Quasars are interesting astrophysical objects that have been recently discovered more comprehensively from the sky surveys. Secondly, some innovative computing algorithms are described about how to classify the light curves of the quasars against light curves of the other stars.

Topics covered in the monograph include Frequency Domain, Correlation, Smoothing, Periodogram, Autoregression, ARIMA Models, Discrimination Analysis, Clustering Analysis, Factor Analysis, Dynamic Fourier Analysis, Random Coefficient Regression, Discrete Fourier Transform, Innovative Computing Algorithms, Knowledge Extraction, Large Complex Databases, Modeling and Simulations, Integration of Hardware, Systems and Networks, Grid Computing, Visualization, Design and Communication, Business Time Series Applications, Biological Time Series Applications, and Astronomical Time Series Applications. The book offers the state of art of tremendous advances in applied time series analysis and innovative computing paradigms and also serves as an excellent reference work for researchers and graduate students working on applied time series analysis and innovative computing paradigms.

The author is grateful for the supports of Dr. Vasile Palade throughout the author's research in Oxford University Computing Laboratory, University of Oxford, UK, and the supports of colleagues in Harvard throughout the author's research in Harvard School of Engineering and Applied Sciences, Harvard University, USA.

May 2009

Sio-Iong Ao, Ph.D.
International Association of Engineers
Unit 1, 1/F, 37-39 Hung To Road
Hong Kong

Contents