Multiobjective Evolutionary Neural Networks for Time Series Forecasting

Swee Chiang Chiam[1], Kay Chen Tan, and Abdullah Al Mamun

[1] Department of Electrical and Computer Engineering
National University of Singapore
4 Engineering Drive 3
Singapore 117576
g0500055@nus.edu.sg

Abstract. This paper will investigate the application of multiobjective evolutionary neural networks in time series forecasting. The proposed algorithmic model considers training and validation accuracy as the objectives to be optimized simultaneously, so as to balance the accuracy and generalization of the evolved neural networks. To improve the overall generalization ability for the set of solutions attained by the multiobjective evolutionary optimizer, a simple algorithm to filter possible outliers, which tend to deteriorate the overall performance, is proposed also. Performance comparison with other existing evolutionary neural networks in several time series problems demonstrates the practicality and viability of the proposed time series forecasting model.

Keywords: Time Series Forecasting, Multiobjective Evolutionary Neural Network.

1 Introduction

Time series forecasting (TSF), the forecast of a chronologically ordered variable, is an important tool in the modeling of complex systems, where the primary aim is to predict the system's behavior without the need to understand its underlying mechanism. The importance of TSF has motivated development in the field of operational research, statistics and computer science for more advanced methodologies and techniques to handle more realistic time series with nonlinear and noisy components.

Neural networks (NN), connectionist models that mimic the central nervous systems, are excellent candidates for TSF due to their capabilities like nonlinear generalization, input-output mapping and noise tolerance. There is currently a wide variety of NN available of different architecture, learning paradigm and etc [1], and studies have shown that NN generally perform better than classical econometric models in TSF. Nevertheless, despite the differences in approach between NN and econometric models i.e. nonlinear generalization versus explicit regression modeling, the ultimate objective in TSF remains in getting the most accurate forecast of the time series i.e. minimizing the error between the forecasted and actual values [2].

S. Obayashi et al. (Eds.): EMO 2007, LNCS 4403, pp. 346–360, 2007.

Designing NN of suitable architecture and connection weights for a given time series is itself a difficult combinatorial optimization problem. This has thus motivated the incorporation of evolutionary optimizers due to their efficacy in dealing with large and complex search spaces. However, the usual approach of evolving NN via training data often results in the over-fitting phenomena, where there is over concentration on the peculiarities of the training set at the expense of losing the regularities needed for good generalization [3]. Consequently, this has inspired the application of evolutionary multiobjective optimization that can simultaneously balance the accuracy and generalization of the evolved NN. The optimization of NN via multiobjective evolutionary optimizers is also known as multiobjective evolutionary neural network (MOENN). Specifically, accuracy and generalization refer to the algorithmic performance of the NN with respect to the training and test data respectively.

As such, this paper investigates the application of MOENN in TSF. While there had been several related works, a more direct approach to balance accuracy and generalization is adopted here, by considering the training and validation accuracy as the objective functions. For this purpose, a TSF model comprising of multiobjective evolutionary algorithm hybridized with particle swarm optimization (PSO) and backpropagation (BP) is proposed. The remainder of the paper is organized as such. Preliminary concepts of TSF and MOENN will be introduced over the next two sections. Following that, experiments to examine the validity of the proposed multiobjective approach will be conducted, before the TSF model is formally presented. Lastly, the performance of the proposed model will be evaluated and compared with other existing algorithms.

2 Time Series Forecasting

Time series is a sequence of observation values of a physical or financial variable ordered at equally spaced time intervals, Δt and is represented as a set of discrete values $x_1, x_2, x_3, x_4, ...etc$. Basically, a TSF model assumes that past patterns will occur in the future and thus, predictions can be done by identifying and generalizing patterns among the past data. As such, the usual training methodology is to split the given time series into two parts, namely the training data, where the learning is performed and the test set, where the performance of the resulting TSF model is measured. The underlying principle is that the TSF model will attune to the training data via the training algorithm and its performance on the test data will approximate its actual performance in real life implementation. The performance in the two separate data is known as training and test accuracy respectively.

Due to the impossibility of deriving a perfect model, there will always be some deviation between the actual and forecasted values and this error, e_t is defined as such,

$$e_t = x_t - \hat{x}_t \tag{1}$$

where $\hat{x}_t$ is the forecasted value for time, t.