An Application of GP-based Prediction Model to Sunspots

Hiroshi YANO*, Ikuo YOSHIHARA*, Makoto NUMATA**,
Tomoo AOYAMA* and Moritoshi YASUNAGA***

* Faculty of Engineering, Miyazaki University Gakuen Kibanadai Nishi, Miyazaki, 889-2192, Japan
(Tel:+81-985-58-7384; Fax:+81-985-58-7384; E-mail:hiroshi@cs.miyazaki-u.ac.jp)
** Electrical and Communication Engineering Dept., Tohoku University Sendai, 980-8579, Japan
(Tel:+81-22-217-7074; Fax: +81-22-217-7074; E-mail:numata@abe.ecei.tohoku.ac.jp)
*** Institute of Info. Sci. & Electronics, University of Tsukuba Tennoudai, Tsukuba, Ibaraki, 305-8573, Japan
(Tel:+81-298-53-5323; Fax:+81-298-53-5206; E-mail:yasunaga@is.tsukuba.ac.jp)

Abstract

We have developed a method to build time series prediction models by Genetic Programming (GP). Our proposed GP includes two new techniques. One is the parameter optimization algorithm, and the other is the new mutation operator. In this paper, the sunspot prediction experiment by our proposed GP was performed. The sunspot prediction is good benchmark, because many researchers have predicted them with various kinds of models. We make three experiments. The first is to compare our proposed method with the conventional methods. The second is to investigate about the relation between a model-building period and prediction precision. In the first and the second experiments, the long-term data of annual sunspots are used. The third is to try the prediction using monthly sunspots. The annual sunspots are a mean of the monthly sunspots. The behaviors of the monthly sunspot cycles in the annual sunspot data become invisible. In the long-term data of the monthly sunspots, the behavior appears and is complicated. We estimate that the monthly sunspot prediction is more difficult than the annual sunspot prediction. The usefulness of our method in time series prediction is verified by these experiments.

1. Introduction

Time series prediction is performed by analyzing the mechanism of the phenomena from past data of time series. Various methods have been proposed for the time series prediction [4, 5, 6, 7]. Most of conventional methods optimize parameters of mathematical models, but to construct the appropriate functional form of the model is difficult. A method for the prediction by Genetic Programming (GP) can construct the functional form of mathematical model. We have proposed a GP-based method to build time series prediction models [1, 2, 3]. Our proposed GP includes two new techniques. One is the parameter optimization algorithm, and the other is the new mutation operator.

This paper uses the sunspot data for the prediction experiment. The sunspot prediction is good benchmark, because many researchers have predicted them with various kinds of models [4, 5, 6, 7]. We make three experiments. The first is to compare our proposed method with the auto-regression (AR) and the standard GP, and to evaluate our proposed method. The second is to investigate about the relation between a model-building period and prediction precision. In the first and the second experiments, the long-term data of annual sunspots are used. The third is to try the monthly sunspot prediction. The annual sunspots are a mean of the monthly sunspots. The behaviors of the monthly sunspot cycles in the annual sunspot data become invisible. In the long-term data of the monthly sunspots, the behavior appears and is complicated. We estimate that the monthly sunspot prediction is more difficult than the annual sunspot prediction. The usefulness of our method is verified by these experiments.

2. Proposed GP

2.1 Time Series Prediction Model

The future value (X_t) is expected by the prediction model with past values (X_{t-1}, X_{t-2},..., X_{t-n}). The model is expressed with the following mathematical function.

$$\tilde{X}_t = f(X_{t-1}, X_{t-2}, \ldots, X_{t-n}) \quad (1)$$

In the GP-based method, the prediction model is described as a tree (Fig.1).

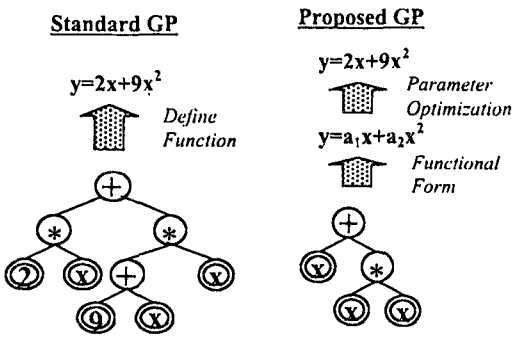

Fig.1 Descriptions of Prediction Model
(Standard GP and Proposed GP)

2.2 Model Building

The building of prediction model is done in two stages. The first is the construction of the functional form of mathematical model by the model production operators, e.g. crossover. In the proposed GP, new mutation operators are took into the model builder. The mutations produce great changes of model structure and make it possible to find better models. The second is the optimization of the model parameter. In the proposed GP, the model parameters are optimized by Back-Propagation-like algorithm. The prediction model is evaluated by the following expression.

$$E = \frac{1}{2}\sum_{i=1}^{k}(\tilde{X}_i - X_i)^2 + C * m \quad (2)$$

$$\text{Fitness} = \frac{1}{E} \quad (3)$$

3. Sunspot Predictions

To evaluate our proposed GP, the prediction experiments using sunspot data were performed. The sunspot prediction is good