A FAST MODEL-BUILDING METHOD FOR TIME SERIES USING GENETIC PROGRAMMING

I. Yoshihara
Faculty of Engineering,
Miyazaki University:
1-1Gakuen-Kibanadai-Nishi,
Miyazaki, 889-2192, Japan
yoshiha@cs.miyazaki-u.ac.jp

T. Aoyama
Faculty of Engineering,
Miyazaki University:
1-1Gakuen-Kibanadai-Nishi,
Miyazaki, 889-2192, Japan
t0b217u@cc.miyazaki-u.ac.jp

M. Yasunaga
Institute of Information Sciences
and Electronics,
University of Tsukuba:
1-1-1,Tennoudai, Tsukuba,
305-8573, Japan
yasunaga@is.tsukuba.ac.jp

INTRODUCTION

To predict time dependent phenomena has a great importance for various fields of real world problems [1]. This paper proposes a fast method to build *one-step-ahead* prediction models with genetic programming (GP) [2]. The method involves two original techniques in crossover and mutation. They are useful for keeping diversity and enable to search rapidly.

GP-BASED TIME SERIES MODEL

One-step-ahead prediction is made by using a set of finite number of past values measured from the system

$$\tilde{x}_t = f(x_{t-1}, x_{t-2}, \ldots, x_{t-n}). \qquad (1)$$

In general, *f*( ) is a complicated function, but can be approximated by composition of arithmetic operators and elementary functions. The functional form of the model is represented by tree-structured chromosome of GP.

(1) Optimization of Model Parameters

After the crossover composes a functional form, the model parameters are optimized by a repetition algorithm imitating *back propagation* of neural networks. The following two steps are performed for each individual in each generation.

step-1 initialization

In each generation, initial values of the parameters in an model (offspring) are inherited from the parents.

step-2 repetition

Each parameter is updated from the root to leaves based on the steepest descent method that is a main part of the BP algorithm, .

The *step-2* is repeated for all the parameters and for all the data until a given terminal condition is satisfied.

(2) Special Mutation

We introduce a special operation in mutation. There are many types of nodes, each of which requires a certain number of branches according to mathematical operations. We consider the number of branches linking to a node from one to four. Usual mutation changes a node merely to another node with the same number of branches, for example, "*cos* --> *sin*", "*mul* --> *add*"(*mul*;multiplication, *add*;addition). Our enhanced mutation, however, change a node to another with different number of branches, for example, "*cos* & *sin* --> *mul*", "*if* --> *add* & *mul*", etc. The following are all the possible combinations upon the number of branches.

case-1	1
case-2	2 , 1 + 1
case-3	3 , 1 + 2 , 2 + 1 , 1 + 1 + 1
case-4	4 , 1 + 3 , 3 + 1 , 2 + 2 , 2 + 1 + 1 , 1 + 2 + 1 , 1 + 1 + 2 , 1 + 1 + 1 + 1

Here *case*-number corresponds to the number of branches connecting to a node. 1 denotes a node with a single branch, 2 with two branches, and so on. Any pairs in the same case are convertible each other *e.g.* "3 --> 2 + 1", "2 + 1 --> 1 + 2"

EXPERIMENTS

We proposed a GP-based model building method for time series and applied it to lots of time series: (1) computer generated chaos *e.g.* Logistic, Loessler, Lorenz, and Mackey-Glass, (2) natural phenomena *e.g.* sunspots, wind velocity, and seismography, and (3) financial problems *e.g.* stock price indices and gold prices.

The experiments lead us to the conclusion that the parameter optimization is always effective but the alternation of nodes is effective only in complex problems.

References

[1] Andreas, S. *et.al.* (ed.): TIME SERIES PREDICTION: FORECASTING THE FUTURE AND UNDERSTANDING THE PAST, Addison-Wesley, 1994

[2] Yoshihara, I., Numata, M., Sugawara, K., Yamaha, S., Abe, K.: "Time Series Prediction Model Building with BP-like Parameter Optimization", CEC-99, 1999