Financial Application of Time Series Prediction based on Genetic Programming

Ikuo YOSHIHARA*, Tomoo AOYAMA* and Moritoshi YASUNAGA**
*Faculty of Engineering, Miyazaki University: 1-1, Gakuen Kibanadai-nishi, Miyazaki, 889-2192, Japan
(Tel:+81-985-58-7384, Fax:+81-985-58-7384, E-mail: yoshiha@cs.miyazaki-u.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 been developing a method to build *one-step-ahead* prediction models for time series using genetic programming (GP). Our model building method consists of two stages. In the first stage, functional forms of the models are inherited from their parent models through crossover operation of GP. In the second stage, the parameters of the newborn model are optimized based on an iterative method just like the *back propagation*.

The proposed method has been applied to various kinds of time series problems. An application to the seismic ground motion was presented in the KACC'99, and since then the method has been improved in many aspects, for example, additions of new node functions, improvements of the node functions, and new exploitations of many kinds of mutation operators. The new ideas and trials enhance the ability to generate effective and complicated models and reduce CPU time.

Today, we will present a couple of financial applications, especially focusing on gold price prediction in Tokyo market.

Keywords: *genetic programming, time series prediction, back propagation, financial problems*

1. Introduction

To build models of time dependent phenomena is of great importance for analysis and prediction in various fields of the real world *e.g.* engineering, industry, business and so forth [3,4,5,9]. The method to build *one-step-ahead* prediction models with genetic programming (GP) has been developed and applied to a plenty of problems *e.g.* computer-generated chaos, natural phenomena, financial problems etc.

Most conventional methods essentially require analysts or designers giving a mathematical model beforehand and optimizing parameters involved in the models. However, to get an idea of appropriate mathematical models is very difficult. On the other hand, the GP-based method does not require giving any model or special knowledge. This is because we use GP, which has succeeded in building models for many kinds of time series [3], for example, chaotic time series [6], system identification and prediction with GP and GMDH [6,7], and prediction function by evolving polynomials [2].

The prediction model is expressed by a composition of numerical operators and elementary functions, and the composite function is represented by a tree-structured graph that corresponds to an individual of GP. Although GP is promising, it is time consuming due to lack of local search ability. Two newly devised techniques are developed to resolve the problem. To enhance searching model functions, we developed parameter optimization method for GP-based model building process, that is, offspring inherit their functional form from their parents, but do not inherit their parameter values from their parents. Parameters are determined by an iterative algorithm similar to the *back propagation* (BP) of the neural networks [1]. We developed sophisticated *mutation* that alters node type, for example, a node with three branches is divided into two nodes with a branch and two branches, or two nodes are put together.

This paper applies the method to financial problems, for instance, stock price, foreign currency exchange, and gold price. Showing overview of financial problems at first, we focus on gold price and investigate a way of application for better prediction. Appendix illustrates an effect of improving algorithm and gives an example of mathematical model generated by GP.

2. GP-based Model-Building Method

2.1 Model Representation

A model is represented by a tree-structured graph, which is a chromosome of GP. The root of the tree is a model function, inner nodes of the tree (branching points) correspond to arithmetic operations or elementary functions, and outer nodes of the tree (leaves) correspond to input variables (table-1).

One-step-ahead prediction is made by using the finite number of past values ($x_{t-1}, x_{t-2}, \ldots, x_{t-n}$) measured from the system:

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

The model function *f*( ) is complicated in general, but assumed to be approximated by composition of arithmetic operators and elementary functions, for example in Table-1. n in Table-1 denotes the number of arguments of the node function or the number of branches of the nodes.