Information Sciences 247 (2013) 62–71

Contents lists available at SciVerse ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

TAIEX forecasting based on fuzzy time series, particle swarm optimization techniques and support vector machines

Shyi-Ming Chen *, Pei-Yuan Kao

Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, ROC

A R T I C L E   I N F O

Article history:
Received 5 August 2012
Received in revised form 11 April 2013
Accepted 5 June 2013
Available online 14 June 2013

Keywords:
Fuzzy sets
Fuzzy time series
Particle swarm optimization
Support vector machines
TAIEX

A B S T R A C T

In this paper, a new method for forecasting the TAIEX is presented based on fuzzy time series, particle swarm optimization techniques and support vector machines. The proposed method to forecast the TAIEX is based on the slope of one-day variation of the TAIEX and the slope of two-days average variation of the TAIEX. Because the slope of two-days average variation of the TAIEX is smoother than the slope of one-day variation of the TAIEX, it is chosen to define the universe of discourse. The particle swarm optimization techniques are used to get optimal intervals in the universe of discourse. The support vector machine is used to classify the training data set. The first feature and the second feature of the support vector machine are the slope of one-day variation and the slope of two-days average variation of the TAIEX, respectively. The experimental results show that the proposed method outperforms the existing methods for forecasting the TAIEX.

1. Introduction

Fuzzy time series has successfully been used to deal with various forecasting problems, such as to forecast the temperature, the economy, the inventory, the earthquake, and the stock index. In [28], Song and Chissom proposed the concepts of fuzzy time series. In [29,30], Song and Chissom presented the time-invariant fuzzy time series model and the time-variant fuzzy time series model to forecast the enrollments of the University of Alabama. In recent years, some fuzzy forecasting methods based on fuzzy time series have been presented [2,4–21,23–27,31–36]. In [2], Chen and Chen presented a method used multivariate to forecast the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) [39]. In [4], Chen presented a high-order fuzzy time series model for forecasting the enrollments of the University of Alabama to get higher forecasting accuracy rates. In [6], Chen and Chen presented a method to forecast the TAIEX by fuzzy variation groups. In [7], Chen and Chung presented a method for forecasting the enrollments of the University of Alabama using fuzzy time series and genetic algorithms. In [8], Chen and Hsu presented a method to forecast the enrollments of the University of Alabama based on fuzzy time series. In [9], Chen and Hwang presented a two-factors time-variant fuzzy time series model to predict the temperature. In [11], Chen et al. presented a method for fuzzy forecasting based on two-factors high-order fuzzy-trend logical relationship groups and particle swarm optimization techniques. In [12], Chen and Wang presented a high-order fuzzy time series forecasting method using fuzzy-trend logical relationships. In [13], Chen and Wang presented a method for forecasting the enrollments of the University of Alabama by using automatic clustering techniques and fuzzy logical relationships. In [15], Huarng observed that the length of intervals in the universe of discourse can affect the forecasting results and proposed the distribution-based length method and the average-based length method for handling the forecasting problems. In [16], Huarng presented a method using a heuristic function to forecast the enrollments of the University of Alabama and the

* Corresponding author. Tel.: +886 2 27376417; fax: +886 2 27301081.
E-mail address: smchen@mail.ntust.edu.tw (S.-M. Chen).

0020-0255/$ - see front matter
http://dx.doi.org/10.1016/j.ins.2013.06.005

Taiwan Futures Exchange (TAIFEX). In [17], Huarng and Yu presented a method using type-2 fuzzy time series to forecast the TAIEX. In [36], Teoh et al. presented a method to forecast the stock markets based on a hybrid multi-order fuzzy time series model, which employed the rough set theory to mine fuzzy logical relationships from time series. In [14], Hung and Lin developed a novel intuitionistic fuzzy least-squares support vector regression with genetic algorithms to forecast the long-term indexes of business cycles. In [26], Lin et al. developed a fuzzy least-squares support vector regression model with genetic algorithms to forecast seasonal revenues. However, because the forecasting accuracy rates of the existing methods are not good enough, a new method is needed to be developed to overcome the drawback of the existing methods to get higher forecasting accuracy rates.

In this paper, a new method for forecasting the TAIEX is presented based on fuzzy time series, particle warm optimization techniques [22] and support vector machines [37]. The proposed method to forecast the TAIEX is based on the slope of one-day variation of the TAIEX and the slope of two-days average variation of the TAIEX. Because the slope of two-days average variation of the TAIEX is smoother than the slope of one-day variation of the TAIEX, it is chosen to define the universe of discourse. The particle swarm optimization techniques are used to get optimal intervals in the universe of discourse. The support vector machine is used to classify the training data set. The first feature and the second feature of the support vector machine are the slope of one-day variation and the slope of two-days average variation of the TAIEX, respectively. The experimental results show that the proposed method outperforms the existing methods for forecasting the TAIEX.

2. Fuzzy time series

In this section, some basic concepts of fuzzy time series are reviewed from [3,28–30], where the values of fuzzy time series are represented by fuzzy sets [38]. Let U be the universe of discourse, where $U = \{u_1, u_2, \ldots, u_n\}$. A fuzzy set A_i in the universe of discourse U can be represented by

$$A_i = f_{A_i}(u_1)/u_1 + f_{A_i}(u_2)/u_2 + \ldots + f_{A_i}(u_n)/u_n, \tag{1}$$

where f_{A_i} is the membership function of the fuzzy set A_i, $f_{A_i}(u_j)$ denotes the degree of membership of u_j belonging to the fuzzy set A_i, $f_{A_i}(u_j) \in [0, 1]$ and $1 \leqslant j \leqslant n$.

Let $Y(t)$ ($t = 0, 1, 2, \ldots$), a subset of R, be the universe of discourse in which fuzzy sets $f_i(t)$ ($i = 1, 2, \ldots$) are defined and let $F(t)$ be a collection of $f_i(t)$ ($i = 1, 2, \ldots$). Then, $F(t)$ is called a fuzzy time series on $Y(t)$ ($t = 0, 1, 2, \ldots$). If there exists a fuzzy relation $R(t-1, t)$, such that $F(t) = F(t-1) \circ R(t, t-1)$, where both $F(t)$ and $F(t-1)$ are fuzzy sets and the symbol "$\circ$" is the max–min composition operator, then $F(t)$ is called derived by $F(t-1)$, denoted by a fuzzy logical relationship shown as follows:

$$F(t-1) \rightarrow F(t).$$

If $F(t-1) = A_i$ and $F(t) = A_j$, where A_i and A_j are fuzzy sets, then the fuzzy logical relationship between $F(t-1)$ and $F(t)$ can be represented by

$$A_i \rightarrow A_j,$$

where A_i and A_j are called the current state and next state of the fuzzy logical relationship, respectively.

3. Particle swarm optimization

In [22], Kennedy and Eberhart developed an optimization algorithm, named particle swarm optimization (PSO), which was inspired by the social behavior of bird flocking or fish schooling. In PSO, a set of particles consists of a particle swarm, where a particle denotes a potential solution. The position and the velocity of the ith particle in an n-dimensional search space can be represented by $X_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,n}]$ and $V_i = [v_{i,1}, v_{i,2}, \ldots, v_{i,n}]$, respectively. Each particle has the personal best position P_{best} which records the position of the best corresponding objective value of the particle. The global position P_{gbest} denotes the position of the best particle in the swarm. The new velocity $V_{i,t}$ and current position X_t of the ith particle is updated as follows:

$$V_{i,t} = \omega \times V_{i,t-1} + C_1 \times r_1 \times (P_{best,i} - X_{i,t-1}) + C_2 \times r_2 \times (P_{gbest} - X_{i,t-1}), \tag{2}$$

$$X_{i,t} = X_{i,t-1} + V_{i,t}, \tag{3}$$

where $V_{i,t}$ denotes the velocity of the ith particle at the tth iteration and $V_{i,t}$ is limited to a pre-defined range $[-V_{MAX}, V_{MAX}]$, the symbol ω denotes the inertial weighting coefficient, C_1 and C_2 are acceleration coefficients, and r_1 and r_2 are two independent random numbers uniformly distributed in the range of [0, 1], respectively. The procedure of a PSO algorithm is shown as follows:

Step 1: Initialize all particles with their respective position and velocity.
Step 2: Update the personal best position $P_{best,i}$ of each ith particle and find the best particle P_{gbest} and its position P_{gbest} in the swarm.
Step 3: Update each particle's velocity and position based on Eqs. (2) and (3), respectively.