Expert Systems with Applications 36 (2009) 7818–7826

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Trading strategy design in financial investment through a turning points prediction scheme

Xiuquan Li, Zhidong Deng *, Jing Luo

State Key Laboratory of Intelligent Technology and Systems, Tsinghua National Laboratory for Information Science and Technology, Department of Computer Science, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Keywords:
Trading strategy
Ensemble artificial neural network
Chaotic analysis
Turning points prediction
Financial time series

ABSTRACT

Turning points prediction has long been a tough task in the field of time series analysis due to its strong nonlinearity, and thus has attracted many research efforts. In this study, the turning points prediction (TPP) framework is presented and further employed to develop a novel trading strategy designing approach to financial investment. The TPP framework is a machine learning-based solution incorporating chaotic dynamic analysis and neural network modeling. It works on the ground of a nonlinear mapping deduced in financial time series through chaotic analysis. An event characterization method is created in TTP framework to characterize trend patterns in ongoing financial time series. The main contributions of this paper are (1) it presents an ensemble learning based TPP framework, within which the nonlinear mapping is approximated by the ensemble artificial neural network (EANN) model with a new parameters learning algorithm; (2) a genetic algorithm (GA) based threshold optimization procedure is described with a newly defined performance measure, named TpMSE, which is used as a cost function; and (3) a trading strategy designing approach is proposed based on the TPP framework. The proposed approach was applied to the two real-world financial time series, i.e., an individual stock quote time series and the Dow Jones Industrial Average (DJIA) index time series. Experimental results show that the proposed approach can help investors make profitable decisions.

1. Introduction

Financial investment activities have become an indispensable component of the everyday life in our modern societies globally. Investors, institutional or individual, are now provided with a tool of financing via multiple channels, i.e., stocks, bonds, commodities, currencies, options and futures, and so forth. However, the high nonlinearity inherent in the behavior of financial markets and the needed professional knowledge make it substantially difficult for investors to make the right investment decisions promptly.

In this context, many efforts have been made to provide decision-making support systems for investors so as to facilitate their analysis work (Chun & Park, 2005; Kodogiannis & Lolis, 2002; Li & Kuo, 2008; Povinelli, 2001; Skabar, 2005; Sun et al., 2005). Among these studies, time series prediction has been proved to be an effective way to help decision making. These studies mainly focused on the issue of next-value prediction, which is forecasting the future value of time series at the oncoming time step, given the historical observation until the current time (Chen, Ji, Zhao, & Nian, 2005; Kodogiannis & Lolis, 2002; Sun et al., 2005). Most of the players in financial markets, however, do not care much about the exact value of next time step; instead, they show great interest in how to predict the future trend of a financial time series and when it is time to alter trading strategy.

Turning points prediction, the prediction of peaks and troughs, can assist us to judge market trend and capture profitable opportunities. It is of great use to both macroeconomic policy-makers and operators in finance world. Some research endeavors trying to tackle this problem have been made with statistical approach (Kling, 1987; Poddig & Huber, 1999; Wecker, 1979). They are mainly developed from the Monte Carlo-based regression approach introduced by Wecker. These methods, based on linear statistical models of ARIMA or VAR, work well with turning points prediction of some macro-economic time series, e.g., a nation's GDP time series. Nevertheless, nonlinear dynamic systems emerge extensively in financial fields. The constraints of stationarity, residual normality, and independence are generally not met for many cases. In fact, turning points are often claimed to be essentially nonlinear phenomena. Nonlinear specifications are better than simpler linear models at reproducing the cycle features of real economic time series, such as GDP (Camacho & Quiros, 2002). Nonlinear models seem to be the natural one to forecast turning points, especially for chaotic time series.

Chaotic time series is usually a sequence of observed values from a complex nonlinear dynamic system with chaotic characteristics.

* Corresponding author. Tel./fax: +86 10 62796830.
E-mail address: michael@tsinghua.edu.cn (Z. Deng).

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.11.014

The existence of chaotic characteristics intrinsic in financial time series, such as in stock market, has been well studied. Previous research has shown the chaotic phenomenon existing in real-world economic and financial time series and proved their chaotic dynamic characteristics (Catherine, Walter, & Michel, 2004; Harrison, Yu, Oxley, Lu, & George, 1999; Hesieh, 1989).

Furthermore, greater success has been achieved in next-value prediction of financial time series by harnessing the chaotic characteristics intrinsic in them (Chen et al., 2005; Jaeger & Haas, 2004). But few efforts have been made on how to predict turning points in financial time series, which is especially important to making trading decisions in financial investment, by exploring the rules of the underlying dynamic system of the chaotic time series.

We hereby propose a machine learning solution based on chaotic analysis and neural network modeling, called TPP framework, for turning points prediction. The theoretical fundamental supporting this proposed method has been provided in our previous work, where a new nonlinear mapping between different data points in primitive time series has been derived and proven by theoretical analysis in reconstructed phase space (Li & Deng, 2007). We have also given the definition of a characteristic function, named turning indicator, to quantitatively represent the event occurrence degree (expression that is used for how likely a point in time series was a turning point) of a turning point and identify the meaningful turning points in time series.

This paper is organized as follows. The main concepts that the TPP scheme has built on are at first described in Sections 2.1 and 2.2. In Section 2.3, the ensemble learning based TPP scheme is introduced. A threshold optimization procedure is presented in Section 2.4, incorporating genetic algorithm with a newly defined performance measure, namely TpMSE, as a cost function. In Section 2.5, we give the parameter learning algorithm for the EANN model. A trading strategy designing approach is brought up in Section 3 based on the TPP scheme. The proposed approach is applied to two real-world financial time series in Section 4, where the experimental results are also given. The performance of the methods is evaluated in Section 5. Finally in Section 6, we draw the conclusion.

2. Turning points prediction scheme

2.1. Theoretical analysis

Our turning points prediction (TPP) scheme is established on the ground of a nonlinear mapping in financial time series deduced through chaotic analysis. The discovered mapping shed light on our model construction process.

The analysis of the dynamic characteristics of one-dimensional time series is traditionally conducted in the reconstructed phase space according to the Taken's theorem. The objective of phase space reconstruction on time series is to rebuild the chaotic attractor in a high dimensional space so as to unveil the dynamic behavior of chaotic time series. Takens has proved that a suitable embedding dimension m can be used to set up a reconstructed space using the delayed coordinate method (Takens, 1981). The reconstructed space is able to resume the tracks of a chaotic attractor when m is properly chosen and $m \geqslant 2d+1$, where d is the dimension of the dynamic system. The phase space reconstruction theory and Taken's theorem provide a solid theoretical basis for the dynamic analysis of chaotic time series.

Our work adopts the Cao method to determine the minimum embedding dimension of chaotic time series (Cao, 1997). It was proved experimentally to be more stable and efficient than the G-P algorithm does (Grassberger & Procaccia, 1983), which has been used in our previous research as well. The Lyapunov exponent quantitatively measures the chaotic characteristic of a chaotic system, named the sensitive dependence on initial condition (SDIC). A lot of experiments have proved that when the trajectories in phase space diverge to a distance e times wider than an initial one, the track is no longer determinable (Huang, 2000; Liu, Zhang, & Yu, 2004). In light of this, the Lyapunov time, also known as the critical predictable interval, determines the short-term predictability of the given chaotic system and is defined as $t_0 = 1/\lambda_1$, where λ_1 is the largest Lyapunov exponent. The numerical computation method proposed by Rosenstein, Collins, and De luca (1993) is employed to compute Lyapunov exponent.

Based on the phase space reconstruction theory and Takens's proof for the differential homomorphism between the reconstructed dynamic system in $\Re^m$ space and the primitive one (Takens, 1981), we discovered that a smooth mapping exists in dynamic evolution of chaotic time series. It further extracts nonlinear dynamic properties of the primitive dynamic system and inspires us to develop a new framework for turning points prediction.

Theorem 1. *Provided that a compact manifold of fractal dimension d in reconstructed $\Re^m$ space is built from chaotic time series through the phase space reconstruction with a properly chosen m, where $m \geqslant 2d+1$, such that the manifold keeps differential homomorphism with the primitive dynamic system, there must be a smooth mapping $\Phi : \Re^{m\tau} \rightarrow \Re^{p\tau}$*

$$[x_{t+1}, x_{t+2}, \ldots, x_{t+p\tau}] = \Phi(x_{t-m\tau+1}, x_{t-m\tau+2}, \ldots x_t), \quad (1)$$

where m is the embedding dimension, τ *the reconstructive delay, and p Lyapunov time.*

The proof for Theorem 1 can be found in Li and Deng (2007).

The derived nonlinear mapping is regarded as an evidence of the existence of certain rules intrinsic in the fluctuation of the values of chaotic time series. It indicates that the rebuilt dynamic system is capable of recalling the track trend within certain time steps by exploring self-similar fractal characteristics of chaotic attractor, which makes it possible to predict turning points of chaotic time series. The TPP framework is thus designed through further research on this nonlinear mapping as described in the following sections. The input layer structure of the neural network model and the computation span of the event characterization function are also expressed in (1). To be more concrete, the input layer of the neural network model should include $m\tau$ neurons and the turning indicator is computed over $p\tau + m\tau$ time steps as described in the next section.

2.2. Event characterization function

A new definition for turning points of time series has been given in our previous work, which leads to a new concept about the criteria for identifying meaningful turning points in financial application, and its reasonableness has also been demonstrated (Li & Deng, 2007).

Definition 1. For a given time series $x_t \in \mathbb{R}$, $\mathbb{t} = \mathbb{1}, \mathbb{2}, \ldots, \mathbb{m}$, a turning point, i.e. peak or trough, is defined as a time step t, such that t is neither located on the upward nor downward side of the time series, and meanwhile, the following value variation (decrease or increase) within $p\tau$ steps exceeds a specific percentage γ.

Having obtained the turning points from the time series defined as in Definition 1, an event characterization function, namely turning indicator, can be defined as a function over the continual $m\tau + p\tau$ time steps, i.e.,

$$\Gamma_\gamma(t) = \Im(x_{t-m\tau+1}, \ldots x_t, \ldots, x_{t+p\tau}) \in [0, 1], \quad (2)$$

$\Gamma_\gamma(t) = 1$ if x_t is a peak while $\Gamma_\gamma(t) = 0$ if x_t a trough. The rest points fall between 0 and 1. The detailed description for the computation method of the turning indicator can be found in Appendix.