A Machine Learning Approach to Predict Turning Points for Chaotic Financial Time Series

Conference Paper · November 2007

DOI: 10.1109/ICTAI.2007.105 · Source: IEEE Xplore

2 authors, including:

Zhidong Deng
Tsinghua University

A Machine Learning Approach to Predict Turning Points for Chaotic Financial Time Series

Xiuquan Li and Zhidong Deng
Department of Computer Science
Tsinghua University, Beijing 100084, China
lixq06@mails.tsinghua.edu.cn

Abstract

In this paper, a novel approach to predict turning points for chaotic financial time series is proposed based on chaotic theory and machine learning. The nonlinear mapping between different data points in primitive time series is derived and proven. Our definition of turning points produces an event characterization function, which can transform the profile of time series to a measure. The RBF neural network is further used as a nonlinear modeler. We discuss the threshold selection and give a procedure for threshold estimation using out-of-sample validation. The proposed approach is applied to the prediction problem of two real-world financial time series. The experimental results validate the effectiveness of our new approach.

1. Introduction

The chaotic characteristic in financial time series has been well studied. Based on the analysis of the chaotic dynamics system underling finance time series, the current research works on time series prediction mainly focus on the problem of next-value prediction [1, 2].

In fact, most of the participants in financial activities, such as portfolio investment, do not care much about the exact value of next time point, instead, they are interested in how to judge the future trend of the financial time series or what is the optimum time for financial operations. Turning points prediction, including peaks and troughs, can assist us with obtaining this information and capturing optimal opportunities. This is a fascinating and challenging problem.

Some research endeavors aiming at tackling this problem have been made by means of statistics [3-5]. They are mainly based on the Monte-Carlo-based regression approach introduced by Wecker. These methods have been applied to the turning points prediction of some macro-economic time series, such as industrial product and GDP of a nation, and attained reasonable performance. Unfortunately, for financial time series, which are inherently noisy, non-stationary, and deterministically chaotic, the constraints of stationarity and residual normality and the independence assumption are generally not met. In fact, turning points are often viewed as nonlinear phenomena. Nonlinear model seems to be the most natural choice to forecast them [6], especially for the financial time series with chaotic characteristics.

In this paper, we propose a novel framework for turning points prediction based on chaotic analysis and artificial intelligence techniques. A nonlinear mapping between different data points in primitive time series is derived in Section 3.1 and proved in Appendix A. A new characteristic function, called turning indicator, is defined in Section 3.2 to quantitatively express the event degree of occurrence of a turning point in time series. The radial basis function (RBF) neural network is used to build a nonlinear model in our framework as described in Section 3.3. Section 3.4 discuss the threshold selection problem and give a procedure for threshold estimation. The experimental results validate the effectiveness of the proposed approach in Section 4. Finally, we draw conclusions.

2. Chaotic analysis of time series

The objective of phase space reconstruction on time series is to unveil the chaotic attractor in a high dimensional space so as to rebulid 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 delay coordinate method [7]. This reconstructed space is able to resume the tracks of a chaotic attractor when m is properly chosen and $m \ge 2d+1$, where d is the dimension of dynamic system. The phase space reconstruction theory and Taken's theorem provides a solid

1082-3409/07 $25.00 © 2007 IEEE
DOI 10.1109/ICTAI.2007.105

theoretical basis for the analysis of the dynamic system in chaotic time series.

The geometry characteristic of chaotic attractor could be studied by analysis of its spatial dimension. Correlation dimension is a kind of fractal dimension and is widely used as an important measure for describing the dynamics characteristic of chaotic attractor. Grassberger and Procaccia presented the G-P algorithm that calculates correlation dimension on the time series [8]. The best embedding dimension for phase space reconstruction can be also calculated using the G-P algorithm.

Another important measure for representing the dynamics characteristic of chaotic system is the Lyapunov exponent. It scores quantitatively a chaotic characteristic of chaotic system, called the sensitive dependence on initial condition (SDIC). A numerical computational method for Lyapunov exponent was proposed by Rosenstein [9].

A lot of experiments have shown that when trajectories diverge to distance e times of initial one, the track is no longer determinable [10,11]. In this case, we have $t_0 = 1/\lambda_1$, where λ_1 denotes the largest Lyapunov exponent, and t_0 is called the Lyapunov time or the critical predictable interval.

3. Turning points prediction approach

3.1. Theoretical analysis

According to phase space reconstruction theory and Takens's proof for the differential homeomorphism between the reconstructed $\Re^m$ space and the primitive dynamic system [7], we can have a homeomorphous dynamic system in a m -dimension phase space. Through the further study of the relationship between the reconstructed dynamic system and the primitive time series, we discover that there exists a smooth mapping that can exhibits more nonlinear dynamics properties of the primitive dynamic system.

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

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

where m is the embedding dimension, τ the embedding delay, and p the largest integer that is less than or equal to Lyapunov time t_0.

The proof for Theorem 1 is given in Appendix A.

Theorem 1 is regarded as an evidence of the existence of some intrinsic rules under the fluctuation of chaotic time series values. 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. The system evolution rules can then be studied. This makes the turning points prediction of chaotic time series possible.

3.2. Event characterization function

According to the definition in [4], a turning point (peak or trough) in the series x_i is defined as any time period t such that x_t is greater (or less) than or equal to the preceding s values and subsequent s values of x_t. Apparently, this needs adaptation for the prediction problem in financial operations. For example, the data point *F* in Fig. 1 complies with the above definition. It, however, is certainly not a turning point in a real sense. Similarly, the data point *C* seems to be a positive trough, but it is not less than all the observations within the s -step range.

In this paper, we give our new definition for turning points. It puts emphasis on the increase or the decrease in the observations of time series after the time period t, instead of the s -step range. According to this definition, the data points *A*, *C*, and *D* in Fig. 1, rather than *B*, *E*, and *F*, can be identified as turning points.

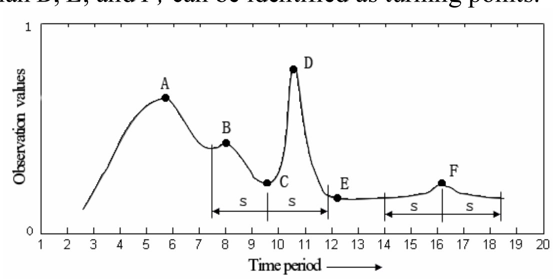

Figure 1. Definition of turning points.

Definition 1. For a given time series $x_t \in \mathbb{R}$, $t = 1, 2, \cdots, n$, a turning point (peak or trough) in the time series x_t is defined as any time period t such that the change (decrease or increase) in the observations of the time series after t exceeds a specific percentage γ within p steps and meanwhile, it is not located on the upward or downward trend of the time series.

To avoid a false turning point that locates on the upward or downward trend of the time series (e.g., the point B in Fig. 1) being identified as a true one, our strategy is to ensure that the type of previously identified turning point is opposite to that of the current one. It means that a peak must follow a valid trough closely and no additional peak is located between them. This ensures the change in the observations on the both sides of a turning point and at the same time, relaxes the restriction on the step range ahead of the turning point.

Based on the turning points obtained from the time series according to the definition 1, an event characterization function, called turning indicator, can be defined as a function over the consecutive $m\tau + p\tau$ time point, i.e.,

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

In the computation of the turning indicator, $\Gamma_{\gamma}(t) = 1$ if x_t is a peak, $\Gamma_{\gamma}(t) = 0$ if x_t a trough, and the other points can be calculated through linear interpolation.

3.3. Nonlinear modeling based on RBF neural network

Considering the nonlinear mapping $\Phi : \Re^{m\tau} \to \Re^{p\tau}$ in (1), it holds that

$$\begin{aligned}&[x_{t-m\tau+1}, \cdots x_t, x_{t+1}, \cdots, x_{t+p\tau}]\\ &= [x_{t-m\tau+1}, \cdots x_t, \Phi(x_{t-m\tau+1}, \cdots x_t)] \\ &= \Phi'(x_{t-m\tau+1}, \cdots x_t)\end{aligned} \quad \textbf{(3)}$$

where Φ' is actually a transformation of Φ.

Combining (3) with (2), we have

$$\begin{aligned}\Gamma_{\gamma}(t) &= \Im(x_{t-m\tau+1}, \cdots x_t, \cdots, x_{t+p\tau}) \\ &= \Im(\Phi'(x_{t-m\tau+1}, x_{t-m\tau+2}, \cdots x_t)) \\ &= \tilde{\Phi}(x_{t-m\tau+1}, x_{t-m\tau+2}, \cdots x_t)\end{aligned} \quad \textbf{(4)}$$

Now the nonlinear mapping $\tilde{\Phi} : \Re^{m\tau} \to \Re^{1}$ is established. The radial basis function network (RBF) is a classical feedforward neural network, which has been applied to a wide range of fields due to a lot of advantages such as fast learning speed, good global optimization, and excellent nonlinear approximation capability. Here we used the RBF neural network to approximate the nonlinear mapping $\tilde{\Phi} : \Re^{m\tau} \to \Re^{1}$.

After the network was well trained based on the training dataset in the reconstructed phase space and the corresponding turning indicators, the network output can then be employed in the prediction phase to represent the probability of occurrence of a turning point in the current time period when the next data points come out.

3.4. Determination of threshold

For the two-class prediction problem, the prediction result depends on how the threshold is selected when the output of the prediction algorithm $\phi(i)$ is continuous. Given a threshold θ, we can define a function $\mathrm{T}(x_i)$ to determine whether a data point x_i is forecasted as a turning point. It follows,

$$\mathrm{T}(x_i) = \begin{cases} 1, & if\, \phi(i) \geq 1-\theta \\ -1, & if\, \phi(i) \leq \theta \\ 0, & otherwise \end{cases}$$

where $\mathrm{T}(x_i) = 1$ indicates a peak and $\mathrm{T}(x_i) = -1$ a trough. $\mathrm{T}(x_i) = 0$ is not a turning point.

In this paper, we give an estimation procedure for the threshold θ based on out-of-sample validation.

After the RBF neural network is well trained using the training dataset, the out-of-sample validation phase including a threshold estimation proceeds. To evaluate the prediction performance, we used the following two measures.

$$PCP = 100\frac{TP}{TP+FN}, \quad PCN = 100\frac{TN}{TN+FP}$$

where PCP and PCN stand for the percentages of correctly predicted positives and correctly predicted negatives, respectively [12].

In the evaluation process, the PCP will increase and the PCN will decrease, both with the increase of threshold θ (as shown in Fig. 2). Thus, there is a trade-off between the PCP and PCN so as to find the optimal threshold θ^*.

In our study, we adopted an intuitional criterion of maximizing the sum of both PCP and PCN to determine the threshold θ^*, i.e., $Q = PCP + PCN$, $\theta^* = \arg\max_{\theta} Q$. The threshold θ^* will be used in the prediction on the testing dataset.

4. Experimental Results

In our experiments, the real-world stock quotes time series of TESCO PLC, which is a FTSE International ingredient stock, was first used here. We used a total of 721 points from Jan. 1, 2004 to Oct. 9, 2006. The whole experimental data was divided into the four parts, including the chaotic analysis data (from Jan. 1, 2004 through Dec.31, 2005), the training data (from Jan. 1, 2006 through May 31, 2006), the validation data (from Jun. 1, 2006 through Jul. 26, 2006), and the testing data (from Jul. 27, 2006 through Oct. 9, 2006). The chaotic analysis results are listed in Tab. 1.

Table 1. Chaotic analysis results of financial time series

Datasets	Embedding Delay	Embedding Dim.	Largest Lya. Exp.	Lya. time
TESCO	1	8	0.1353	7.3931
DJIA	1	8	0.1476	6.7767

The RBF neural network was built as described in Section 3.3 and trained using the training dataset. Through our threshold estimation procedure, a threshold of 0.2 was computed based on out-sample validation, as shown in Fig.2. (a).

In our second experiment, similar chaotic analysis was conducted on Dow Jones Industrial Average (DJIA) time series taken from Sep. 10, 2001 to Nov. 16, 2004. We trained RBF neural network on the DJIA time series from Nov. 17, 2004 to Jul. 8, 2005. Then the validation process and the threshold determination were conducted on the dataset from Jul. 9, 2005 to Nov. 14, 2005. In this case, we had a threshold of 0.23 as shown in Fig.2 (b).

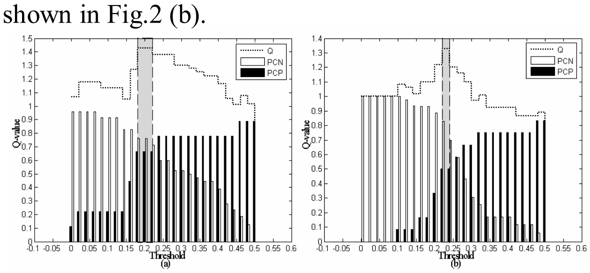

Figure 2. Determination of threshold.

Finally, the dataset from Nov. 15, 2005 to Feb. 22, 2006 was employed as the testing dataset to validate the prediction performance. The prediction results are shown in Fig.3(a) (TESCO) and Fig.3(b) (DJIA), respectively. When the prediction of turning indicator exceeds the thresholds indicated by the two horizontal lines in the lower subfigures, the time point is predicted as a turning point, which is highlighted by shadow bars (light gray for peak and dark gray for trough).

5. Conclusion

In this paper, we propose a novel machine learning approach to predict turning points of chaotic time series. The nonlinear mapping between different data points in primitive time series is derived and a characteristic function is defined. We then establish the framework of turning points prediction. We also discuss the problem of threshold selection and give a procedure based on out-of-sample validation. The experimental results on real-world financial time series show that our approach has good performance. Although it cannot reach a high accuracy in the exact value prediction due to the sensitive dependence on initial condition (SDIC) and data noise, we can still use the characteristic of short-term memory to predict the trend or the curve shape of the next several time steps, which can be further used in the event characterization function for turning points prediction. Apparently, our new approach provides an assistant tool for financial researchers and participants in financial activities.

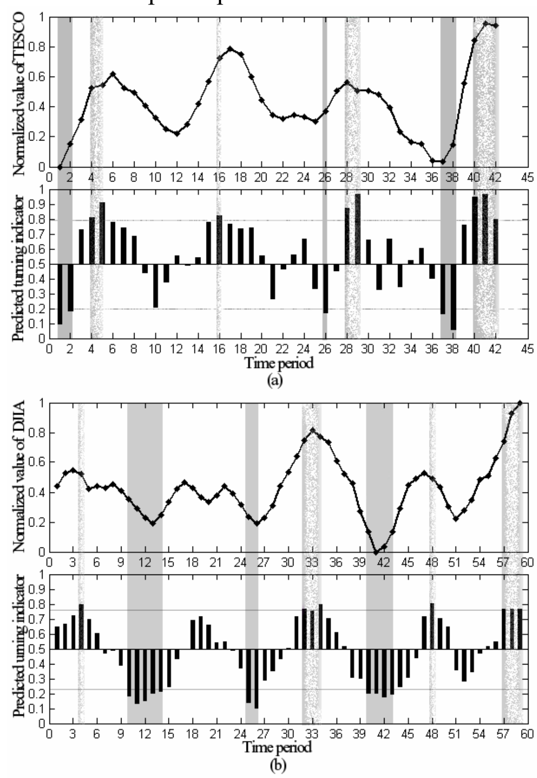

Figure 3. Prediction results on the testing dataset.

Appendix A.

Proof of Theorem 1. According to Taken's theorem, we can get a compact manifold of fractal dimension d in m -dimension phase space based on phase space reconstruction, where $m \geq 2d+1$ [10]. The reconstructed dynamic system in $\Re^m$ space keeps differential homeomorphism with the primitive dynamic system with one observable, which means that there exists a smooth mapping $\varphi : \Re^m \to \Re^m$ that describes nonlinear dynamics of primitive dynamic system, i.e.,

$$y_{t+\tau} = \varphi(y_t)$$

where τ is an embedding delay time.

Let p denote the largest integer that is less than or equal to Lyapunov time t_0. There exist mappings $\varphi, \varphi^2, \varphi^3, \cdots, \varphi^p$ in m -dimensional metric space

$$\begin{aligned} y_{t+\tau} &= \varphi(y_t), \\ y_{t+2\tau} &= \varphi^2(y_t), \\ &\vdots \\ y_{t+p\tau} &= \varphi^p(y_t) \end{aligned} \quad \textbf{(5)}$$

Note that every point in $\Re^m$ space can be mapped to a real vector of m length in primitive time series through time-delay embedding. It follows

$$\begin{cases} y_t = [x_{t-(m-1)\tau}, x_{t-(m-2)\tau}, \cdots, x_t] \\ \vdots \\ y_{t+p\tau} = [x_{t-(m-p-1)\tau}, x_{t-(m-p-2)\tau} \cdots, x_{t+p\tau}] \end{cases} \quad \textbf{(6)}$$

where m is an embedding dimension.

Replacing variables in (5) by (6), we have

$$\begin{cases} [x_{t-(m-2)\tau}, \cdots, x_{t+\tau}] = \varphi(x_{t-(m-1)\tau}, \cdots, x_t), \\ \vdots \\ [x_{t-(m-p-1)\tau}, \cdots, x_{t+p\tau}] = \varphi^p(x_{t-(m-1)\tau}, \cdots, x_t) \end{cases} \quad \textbf{(7)}$$

Since each item on the left side of (7), except for the last one, also appears on the right side of it, (7) can be rewritten as

$$\begin{cases} x_{t+\tau} = \varphi'(x_{t-(m-1)\tau}, \cdots, x_t), \\ \vdots \\ x_{t+p\tau} = \varphi'^p(x_{t-(m-1)\tau}, \cdots, x_t) \end{cases} \quad \textbf{(8)}$$

Thus, it is easy to have

$$(x_{t+\tau}, \cdots, x_{t+p\tau}) = \Phi_1(x_{t-(m-1)\tau}, x_{t-(m-2)\tau}, \cdots, x_t) \quad \textbf{(9)}$$

Similarly, the nonlinear mappings can be derived by substituting the index t with $t-1, \cdots, t+1-\tau$, respectively, that is

$$\begin{cases} (x_{t-1+\tau}, \cdots, x_{t-1+p\tau}) = \Phi_2(x_{t-1-(m-1)\tau}, \cdots, x_{t-1}) \\ \vdots \\ (x_{t+1}, \cdots, x_{t+1+(p-1)\tau}) = \Phi_\tau(x_{t-m\tau+1}, \cdots, x_{t+1-\tau}) \end{cases} \quad \textbf{(10)}$$

Finally, the nonlinear mapping between different data points in primitive time series is derived by algebra combination as follows,

$$[x_{t+1}, x_{t+2}, \cdots, x_{t+p\tau}] = \Phi(x_{t-m\tau+1}, x_{t-m\tau+2}, \cdots x_t) \quad \textbf{(11)}$$

Acknowledgements

This work was supported in part by the National Science Foundation of China under Grant No. 60621062.

References

[1] H. Jaeger and H. Haas, "Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication", *Science*, vol.304, 2004, pp. 78-80.

[2] F. Chen, G.R. Ji, W.C. Zhao, and R. Nian, "The Prediction of the Financial Time Series Based on Correlation Dimension", *Lecture Notes in Computer Science*, vol.3610, 2005, pp. 1256-1265.

[3] F.W. Wecker, "Predicting the Turning Points of a Time Series", *Journal of Business*, vol.52, 1979, pp. 35-50.

[4] J.L. Kling, "Predicting the Turning Points of Business and Economic Time Series", *Journal of Business*, vol.60 no.2, 1987, pp. 201-238.

[5] T. Poddig and C. Huber, "A Comparison of Model Selection Procedures for Predicting Turning Points in Financial Time Series", *PKDD'99, LNAI*, vol.1704, 1999, pp. 492-497.

[6] M. Camacho and G. Perez-Quiros, "This is What the Leading Indicators Lead", *Journal of Applied Econometrics*, vol.17, 2002, pp. 61–80.

[7] F. Takens, "Detecting Strange Attractors in Turbulence", *Lecture notes in Math*, vol.898, 1981, pp. 361-381.

[8] P. Grassberger and I. Procaccia, "Measuring the Strangeness of Strange Attractors", *Physica D*, vol.9, 1983, pp. 189-208.

[9] M.T. Rosenstein, J.J. Collins, and C.J. De luca, "A Practical Method for Calculating Largest Lyapunov Exponents from Small Data Sets", *Physica D*, vol.65, 1993, pp. 117-134.

[10] R.S. Huang, *Chaos and Its Application*, Wuhan University Press, Wuhan, 2000, pp. 115-135.

[11] J.Q. Liu, T.Q. Zhang, and S.K. Yu, "Chaotic Phenomenon and the Maximum Predictable Time Scale of Observation Series of Urban Hourly Water Consumption", *Journal of Zhejiang University*, vol.5, no.9, 2004, pp. 1053-1059.

[12] Pierre. Baldi, Soren. Brunak, Yves. Chauvin, Claus A.F. Andersen, and Henrik. Nielsen, "Assessing the Accuracy of Prediction Algorithms for Classification: An Overview", *Bioinformatics*, vol.16, no.5, 2000, pp. 412-424.