Expert Systems with Applications 38 (2011) 11347–11354

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Intelligent stock trading system based on improved technical analysis and Echo State Network

Xiaowei Lin, Zehong Yang *, Yixu Song

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

ARTICLE INFO

Keywords:
Stock trading system
Technical analysis
Genetic Algorithm
Echo State Network

ABSTRACT

Stock trading system to assist decision-making is an emerging research area and has great commercial potentials. Successful trading operations should occur near the reversal points of price trends. Traditional technical analysis, which usually appears as various trading rules, does aim to look for peaks and bottoms of trends and is widely used in stock market. Unfortunately, it is not convenient to directly apply technical analysis since it depends on person's experience to select appropriate rules for individual share. In this paper, we enhance conventional technical analysis with Genetic Algorithms by learning trading rules from history for individual stock and then combine different rules together with Echo State Network to provide trading suggestions. Numerous experiments on S&P 500 components demonstrate that whether in bull or bear market, our system significantly outperforms buy-and-hold strategy. Especially in bear market where S&P 500 index declines a lot, our system still profits.

1. Introduction

Data mining in stock market has been a hot topic for a long time due to its potential profits. Unfortunately, stock market is a complex and dynamic system with noisy, non-stationary and chaotic data series (Peters, 1994). Stock movement is affected by complicated factors, which can be divided into two groups: one is determinant, such as gradual power change between buying and selling side; the other is random factors, such as emergent affairs or daily operation variations (Bao & Yang, 2008). Therefore, data mining in stock market is very difficult and challenging. Recently, advances in artificial intelligence have led to a number of interesting new approaches to stock data mining, based on non-linear and non-stationary models. Among them, soft computing techniques, such as fuzzy logic, neural networks and probabilistic reasoning draw most attention because of their ability to handle uncertainty and noise in stock market (Vanstone & Tan, 2003, 2005). Applications range from time series prediction, classification to rule induction.

Although past studies have attained remarkable achievement in stock data mining, especially price prediction, they seldom directly guide trading. Future price forecast is not enough to suggest ideal trading operation to get profit as much as possible. An ideal trading operation should occur at the peak or bottom of price trend, that is, a good investor will sell stocks near the top of the trend and buy them close to the bottom. Thus, it is important to predict not only the future price but also when the price trend will hit the peak or bottom. In real market, technical analysis is widely used to assist decision-making. Its central idea is to look for peaks, bottoms, trends and indicators to estimate the possibility of current trend reversal and then make buy/sell decisions based on technical indicators which are some statistics derived from recent historical data (Bao & Yang, 2008). However, traditional technical analysis suffers from some shortcomings. First, it is difficult to directly apply technical analysis on individual stocks, especially for green hand. Technical analysis usually appears in a form as a trading rule. Take the popular "Golden Cross" and "Dead Cross" for example, if the sigh of (long-term moving average) – (short-term moving average) changes from positive to negative, it is called "Golden Cross" which indicates to buy stocks; if the sign of (long-term moving average) – (short-term moving average) changes from negative to positive, it is called "Dead Cross" which suggests to sell stocks. In the above description, it is hard to decide the time spans for both long-term and short-term moving average (MA) because each stock should have its own appropriate time spans. Investors usually choose those parameters according to their experience. Second, there are various technical analysis approaches, such as moving average approach, relative strength indicator (RSI) approach and stochastic indicator approach. Not all of them are effective for every stock. How to choose proper technical analysis methods for individual stock is also difficult for ordinary investors.

In this paper, we propose an intelligent stock trading system based on enhanced technical analysis and neural network. Genetic

* Corresponding author. Tel.: +86 10 62796828; fax: +86 10 62782266.
E-mail addresses: xiaow.lin@gmail.com (X. Lin), yangzehong@sina.com (Z. Yang), songyixu@sohu.com (Y. Song).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.03.001

Algorithm (GA) is utilized to improve traditional technical analysis by learning appropriate parameters for each trading rule. Then, the improved trading rules behave as experts together to give trading suggestions with a novel neural network-Echo State Network (ESN). The experiments demonstrate that whether in bull or bear market, our system will gain more income than buy-and-hold strategy. Particularly, it can still earn in bear market.

The rest of the paper is organized as follows: Section 2 describes the application of GA to improve traditional technical analysis; Section 3 introduces ESN and our system; Section 4 shows the experiments and results. Finally, we make a conclusion and suggest for further research.

2. Technical analysis enhancements

Technical analysis tends to forecast future price movements based on the study of past markets. It assumes that history will repeat itself and tries to identify archetypal patterns which have appeared in the past to predict what is likely to happen in the future. Although it has been recognized as one of the most reliable techniques for dealing stocks (Baba, Kawachi, Nomura, & Sakatani, 2004), it is not convenient to utilize technical analysis directly because it often appears as trading rules with parameters which have to be determined through experience. In this section, we improve traditional technical analysis with GA.

2.1. Genetic Algorithm

GAs are heuristic search techniques that are based on the theory of natural selection and evolution (Holland, 1992). They are particularly suitable for multi-parameter optimization problems in which an object function is subject to numerous hard and soft constraints (Kim, Min, & Han, 2006; Kim & Shin, 2007). In this paper, GA helps to enhance traditional technical analysis by generating a combination of parameters with which the corresponding trading rule will identify optimal trading points as close as possible to real reversal points of trends.

GA usually consists of four stages: initialization, selection, crossover and mutation. In the initialization stage, a population of genetic structures, called chromosomes that are randomly distributed in the solution space, is selected as the starting point of the search (Kim & Shin, 2007). Then, each chromosome, which represents a potential solution of the target problem, is evaluated by a user-defined fitness function. Through selection, the chromosomes with high performance will be preserved and propagate from generation to generation. The crossover forms a new offspring between two randomly selected "good parent" (Kim & Shin, 2007). And the mutation guarantees that it is possible to reach any point in the search space.

For real-world applications of optimization problems, choosing fitness function is the most critical step (Kim & Shin, 2007). In this paper, we design the fitness function to measure how close the suggested trading points are to those turning points of price trends. Suppose that there is an expected trading point sequence $T = \{T_1, T_2, \ldots, T_n\}$, in which buying and selling signals are staggered. For every expected trading point T_i, we search for operation signal S_j given by a specified technical analysis approach between its last and the next expected trading point ($T_{i-1} < S_j < T_{i+1}$).

(1) If T_i is an expected buying point, there are three cases:

(a) If S_j is a suggested buying point, the value of fitness function at T_i is computed as follows:

$$fitness(T_i) = close(S_j) - close(T_i) \quad (1)$$

in which $close(S_j)$ and $close(T_j)$ represent closing prices at S_j and T_i, respectively. The closer the closing price of S_j is to that of T_i, the smaller the value of fitness function is.

(b) If S_j is a suggested selling point and its price is close to T_i (satisfying $|close(S_j)\text{-}close(T_i)|/close(T_i) < 0.05$), then the value of the fitness function at T_i is

$$fitness(T_i) = 2 \times (\max(close(T_{i-1} : T_{i+1})) - close(T_i)) \quad (2)$$

in which $\max(close(T_{i-1}{:}T_i))$ means the maximum closing price between T_{i-1} and T_{i+1}. Since it is wrong to sell near the bottom of trend, such operation will be punished seriously.

(c) If no suggested operations are found between $T_{i-1}+1$ and $T_{i+1}-1$, it should also be punished due to missing opportunity. The fitness function is

$$fitness(T_i) = \max(close(T_{i-1}+1 : T_{i+1}-1) - close(T_i) \quad (3)$$

(2) Similarly, if T_i is an expected selling point, the fitness function is designed as follows:

(a) If S_j is a suggested selling signal, the fitness function at T_i is

$$fitness(T_i) = close(T_i) - close(S_j) \quad (4)$$

(b) If the price at S_j is close to that at T_i and S_j is misjudged as a buying point, the fitness function is

$$fitness(T_i) = 2 \times (close(T_i) - \min(close(T_{i-1} : T_{i+1})) \quad (5)$$

in which $\min(close(T_{i-1}{:}T_i))$ means the minimum closing price between T_{i-1} and T_{i+1}.

(c) If trading opportunity has been missed between $T_{i-1}+1$ and $T_{i+1}-1$, the fitness function is

$$fitness(T_i) = close(T_i) - \min(close(T_{i-1}+1 : T_{i+1}-1) \quad (6)$$

Finally, the fitness function of a suggested trading series $S = \{S_1, S_2, \ldots, S_m\}$ is defined as follows:

$$fitness(S) = \sum_{i=1}^{n} fitness(T_i) \quad (7)$$

GA is expected to find an optimal combination of parameters which will make the value of the fitness function minimum.

2.2. Moving average system

Simple MA is a popular technical indicator which calculates the mean price in a specified period. There are several methods to apply MA. Here, we improve two of them: one is the famous "Golden Cross" and "Dead Cross"; the other is MA envelope approach.

2.2.1. "Golden Cross" and "Dead Cross"

Baba et al. (2004) proposes a method to detect "Golden Cross" and "Dead Cross" with GA to reduce chance losses, which is also adopted in our system. Assume $Z_t = MA(N) - MA(n)$, in which $MA(N)$ means long-term MA while $MA(n)$ means short-term MA.

(a) If $z_t \geqslant 0$, find a number t_1 which is the closest to t ($t_1 < t$), and satisfies $z_{t_1-1} \leqslant 0$ and $z_{t_1} > 0$. Let $\overline{Mz_t} = \max(z_{t_1}, z_{t_1+1}, \ldots, z_t)$, if conditions (8) and (9) are met, it is time to buy

$$\overline{Mz_t} > b \times c \quad (8)$$

$$z_t < \min(\overline{Mz_t}/a, c) \quad (9)$$

Note that a, b, c are parameters.

(b) If $z_t < 0$, find the number k_1 which is closest to k ($k_1 < k$), and satisfies $z_{k_1-1} \geqslant 0$ and $z_{k_1} < 0$. Let $w_k = -z_k (k = k_1, \ldots, k)$, $\overline{Mw_k} = \max(w_{k_1}, w_{k_1+1}, \ldots, w_k)$, if conditions (10) and (11) are met, it is time to sell.