Knowledge-Based Systems 66 (2014) 68–81

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

A novel approach to dynamic portfolio trading system using multitree genetic programming

Somayeh Mousavi, Akbar Esfahanipour *, Mohammad Hossein Fazel Zarandi

Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran

ARTICLE INFO

Article history:
Received 4 November 2013
Received in revised form 9 April 2014
Accepted 10 April 2014
Available online 2 May 2014

Keywords:
Dynamic portfolio trading system
Multitree genetic programming
Technical indices
Trading rules
Tehran stock exchange
Toronto stock exchange

ABSTRACT

Dynamic portfolio trading system is used to allocate one's capital to a number of securities through time in a way to maximize the portfolio return and to minimize the portfolio risk. Genetic programming (GP) as an artificial intelligence technique has been used successfully in the financial field, especially for the forecasting tasks in the financial markets. In this paper, GP is used to develop a dynamic portfolio trading system to capture dynamics of stock market prices through time. The proposed approach takes an integrated view on multiple stocks when the GP evolves and generates a rule base for dynamic portfolio trading based on the technical indices. In the present research, a multitree GP forest has been developed to extend the GP structure to extract multiple trading rules from historical data. Furthermore, the consequent part of each trading rule includes a function rather than a constant value. Besides, the transaction cost of trading which plays an important role in the profitability of a dynamic portfolio trading system is taken into account. This model was used to develop dynamic portfolio trading systems. The profitability of the model was examined for both the emerging and the mature markets. The numerical results show that the proposed model significantly outperforms other traditional models of dynamic and static portfolio selection in terms of the portfolio return and risk adjusted return.

1. Introduction

Dynamic portfolio optimization consists of portfolio selection problem in which we find the optimum way of investing a particular amount of money in a given set of securities through time. In this problem, we have to answer three questions, which assets are proposed to buy/sell, how much and when. This definition of portfolio selection problem regards to active portfolio management approach. An alternative approach is passive portfolio management, in which the investor establishes a well diversified portfolio and maintains it for a specific period of time. The portfolio selection problem with the passive strategy is also referred as static portfolio selection.

The foundation of static portfolio selection was laid by Markowitz in 1952 [1]. He proposed a mean–variance optimization model to design an optimum portfolio based on the idea of minimizing the risk and maximizing the expected returns. After the Markowitz's mean–variance model, many other models use its fundamental assumptions [2]. In all these models, today known as the classic models, the portfolio expected return is given by the linear combination of the expected returns of the participating stocks in the portfolio. However, the literature reports many portfolio risk measures which are often based on the moments of returns of the participating stocks. These models can be viewed as an extension of Markowitz model with one, two or more objectives including some constraints such as minimum transaction lots, cardinality and boundary constraints. In order to solve the portfolio selection problem, many studies utilized artificial intelligence (AI) techniques such as the artificial neural networks [3,4] as well as the metaheuristics like simulated annealing, Tabu search and genetic algorithms [5,6].

Alternatively, the active strategy tries to find undervalued or overvalued stocks dynamically in order to achieve a significant profit. One approach is to use the technical analysis to predict stock price trends. The earlier studies have utilized the linear methods of time series analysis to forecast the stock price trends [7]. However, the linear models are established based on some assumptions such as linearity and normal distribution of stock prices. Since the stock market is a highly nonlinear dynamic system, the optimum portfolio is rather difficult to construct through the mathematical way only. Therefore, AI techniques look more promising since they have the ability to deal with the complex nonlinear problems and

* Corresponding author. Address: No 24, Hafez Ave, P.O. Box: 15875-4413, Tehran, Iran. Tel.: +98 (21) 6454 5369; fax: +98 (21) 6695 4569.

E-mail address: esfahaa@aut.ac.ir (A. Esfahanipour).

http://dx.doi.org/10.1016/j.knosys.2014.04.018
0950-7051/

they are self-adaptive for dynamically changing problems. Among these techniques, artificial neural networks, genetic algorithms and genetic programming are the most applied in technical financial forecasting [8].

As a main approach in the AI field, artificial neural network (ANN) has been widely used because of its ability to forecast financial instruments [9–13]. A vast literature about ANN applications in financial forecasting and trading system development has been presented in [14]. In this study, ANN was employed to propose a prediction-based portfolio optimization model that can capture short-term investment opportunities. Despite the wide spread use of the artificial neural networks in the financial domain, there are significant problems that must be addressed. ANNs are data driven models and the underlying rules in the data are not always apparent, which leads to so called black box models. Therefore, the investors cannot benefit from the knowledge discovery in the analytic process.

Genetic algorithm (GA) is another AI technique which has been applied successfully to financial problems. As one of the most popular heuristic optimization techniques, GA was originally developed by Holland in 1975 [15]. This search technique has been widely used because it is simple and has no restrictive assumptions about the solution space. Oh et al. used genetic algorithms to support portfolio optimization for index fund management [16]. In their later work, they proposed a new portfolio selection algorithm based on the portfolio beta using GA [17]. Lin and Liu extended Markowitz model with minimum transaction lots and used GA as their solver [18]. Moreover, GA is utilized to solve the classical portfolio optimization of Markowitz with minimum transaction lots, cardinality constraints and sector capitalization [19]. Although GA has been successfully applied to solving portfolio optimization problems as well as ANN [3,4], they showed a good performance for static portfolio selection with passive strategy. In the context of trading system development, GA was used to exploit technical trading rules in the US exchange market [20]. These trading rules led to positive excess returns in comparison with buy and hold strategy. In [21], a genetic algorithm was used to find technical trading rules for Standard and Poor's Composite Stock Index. In [22], a GA model was proposed to build an associative classifier that can discover trading rules from many numerical technical indicators. In this study, the associative classification rules were extracted to express relations between numerical data, for the first time. In this structure, the left side of the associative classification rules contains a set of trading signals, advised by the technical indicators, and the right side indicates buying or selling signals. In another research [23], a hybrid model integrating GA and support vector machines (SVM) was proposed to explore stock market trends. In this model, GA was applied to selecting the appropriate combination of input features and parameters of SVM and then least squares SVM was evolved to predict stock market movement direction in terms of historical data series. Integration of GA and ANNs was also applied successfully to predict stock markets [24,25]. In [24], the self organizing map neural network was used to cluster the data set and the GA was used to extract a fuzzy rule base from historical data. In [25], GA was used as a global search method to evolve neural networks initial weights and then the Levenberg–Marquardt back propagation algorithm was used as a local search method to tune the obtained weights.

However, genetic programming (GP) seems to be more appropriate in extracting knowledge from data, because of its structure. GP is viewed as the extension of GA, developed by Koza in 1992 [26]. The main difference between these two approaches is in the representation scheme used. GA uses string representations which prepare solutions for the problem, whereas GP represents individuals as executable programs which can solve the problem. So far GP has been applied successfully to a wide range of financial fields such as portfolio insurance [27], bankruptcy prediction [28], nonlinear time series forecasting [29] and stock trading systems [30–35]. In [27], a dynamic proportion portfolio insurance strategy was extended to generate a time adaptive risk multiplier according to the market conditions. In this strategy, GP is used to evolve an equation tree for the risk multiplier based on the risk variables. Also, GP with decision tree structure is utilized to classify bankrupt and non-bankrupt firms based on the firm's financial ratios [28]. Moreover, the ability of GP in forecasting the nonlinear time series is confirmed. In [29] a hybrid forecasting model was proposed to forecast nonlinear time series by combining autoregressive integrated moving average (ARIMA) with GP. In this approach, the ARIMA was utilized to model the linear component of time series and the GP was utilized to model the nonlinear component of time series. This study focused on the GP's ability to derive a mathematical equation from small data sets comparing with ANNs and SVM. In GP approach for trading system development, trading rules are extracted from historical data in the form of IF–THEN rules. In [30–33], the trading rules were represented as decision trees with binary or trinary outputs. The mentioned studies show that GP could generate profitable binary technical trading rules with buy and sell signals in risk-unadjusted basis [31] as well as trinary trading rules with buy, hold and sell signals when risk and transaction cost are considered [33]. Chen et al. [34,35] have used genetic network programming (GNP) to develop stock trading systems. GNP is an extension of GP in terms of individuals with graph structures. Recently, Mabu et al. [36] have extended a GNP with rule accumulation (GNP-RA) algorithm for decision making in stock trading. In GNP-RA, a large number of rules are extracted throughout the generations instead of one rule. They showed the rule-based stock trading model outperforms the conventional individual-based stock trading model.

All of the mentioned studies have developed dynamic trading systems for tracking individual securities or market indices. However, it is proved that the risk of investment can be reduced by diversification, i.e. investing in a portfolio of securities rather than one security. Hence, the most of investors tend to use portfolio selection and management tools. In the context of portfolio selection, the above-mentioned studies tried to select an optimum portfolio using passive portfolio management approach. They did not utilize technical analysis indices and technical trading rules to develop a dynamic portfolio trading system.

Ghandar and his colleagues extended a fuzzy rule based system for dynamic portfolio trading using evolutionary algorithms [37]. Their portfolio trading system uses some technical indicators of the stocks as the inputs, to rank the stocks according to their fitness function as a "buy" recommendation. Notwithstanding the good performance of their system which is accepted by both the financial industry and academia, there are three ways to improve their model from our point of view. First, Ghandar and his colleagues have provided an identical rule base for all stocks, whereas each stock may fluctuate in a different pattern. Therefore, the performance of the system can be improved by recognizing a specific pattern which can be represented by a unique set of rules for each stock. Second, the cash earned by selling the worst-ranked stocks is distributed evenly over the best-ranked stocks, whereas this distribution is not optimal in the context of portfolio optimization. In other words, the changes in the stocks weights are not determined by their system and the system uses the same weights to include the best-ranked stocks in the case of portfolio rebalancing. Third, their system uses well-known technical indicators as input variables, for example 20 days moving average (MA) with lag of 20 days. However, the optimal parameters of technical indicators are case specific and should be determined accordingly. For example, the selected lag can influence the profitability of the MA rule [38].