Information Sciences 346–347 (2016) 261–274

Contents lists available at ScienceDirect

Information Sciences

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

An intelligent pattern recognition model for supporting investment decisions in stock market

Tai-liang Chen[a,*], Feng-yu Chen[b]

[a] Department of Digital Content Application and Management, Wenzao Ursuline College of Languages, 900 Mintsu, 1st Road, Kaohsiung 807, Taiwan, ROC

[b] Computer and Network Center, National Kaohsiung University of Applied Science, Chien Kung Campus 415 Chien Kung Road, Kaohsiung 807, Taiwan, ROC

ARTICLE INFO

Article history:
Received 10 December 2012
Revised 2 April 2015
Accepted 29 January 2016
Available online 10 February 2016

Keywords:
Stock market forecasting
Perceptually important point (PIP) identification matching
Template matching technique
Chart pattern
Technical indicator

ABSTRACT

For many years, how to make stock market predictions has been a prevalent research topic. To carry out accurate forecasting, stock analysts and academic researchers have tried various analysis techniques, algorithms, and models. For example, "technical analysis" is a popular approach used by common stock investors to analyze market trend, and Artificial Intelligence (AI) algorithms such as genetic algorithms (GAs), neural network (NN), and fuzzy time-series (FTS), were proposed by researchers to forecast the future stock index. Although the daily forecasts are very useful for professional investors who implement intraday trading, we argue that forecasting a bullish turning point is a more interesting issue than the future stock index for common investor because an accurate forecast will bring a huge amount of stock return. Therefore, this paper proposes an intelligent pattern recognition model, based on two new stock pattern recognition methods, "PIP bull-flag pattern matching" and the "floating-weighted bull-flag template," to recognize a bull-flag stock pattern. The bull-flag pattern is a stock's turning point with proper timing, which can enable a stock investor to profit. To promote recognition accuracy, the proposed model employs chart patterns and technical indicators, simultaneously, as pattern recognition factors. In the model verification, we evaluate the proposed model with stock returns by forecasting two stock databases (TAIEX and NASDAQ), and comparing the returns with other advanced algorithms. The experimental results indicate that the proposed model outperforms the published algorithms, such as rough set theory (RST), genetic algorithms (GAs) and their hybrid model, and gives a high-level of profitability. Additionally, the trading strategies, provided by the proposed model, also help investors to make beneficial investment decisions in the stock market.

1. Introduction

Stock market is a very complex and changeable system influenced by many factors such as economic environment, political policy, industrial development, and market news, etc. To make profit from the capital market, stock investor has been looking for right tools and techniques to analyze stock market. As we known, the technical analysis [9] is a popular approach used by investment professionals and common investors. The principles of technical analysis are derived from the

* Corresponding author: Tel.: +886 (0)920975168; fax: +886 (0)73105785.
E-mail address: 97007@mail.wzu.edu.tw, too.cool.chen@gmail.com (T.-l. Chen).

http://dx.doi.org/10.1016/j.ins.2016.01.079
0020-0255/

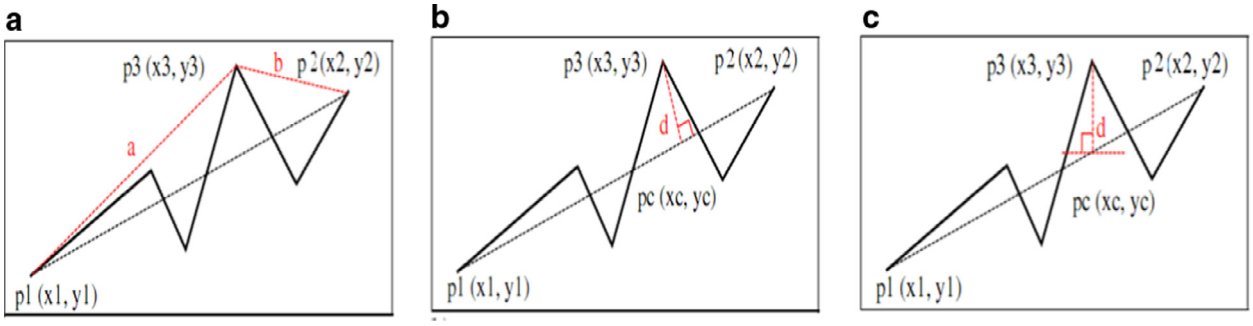

Fig. 1. Distance measure methods for PIP identification: (a) Euclidean distance based: PIP-ED, (b) perpendicular distance based: PIP-PD and (c) vertical distance based: PIP-VD [11].

observations of stock markets over hundreds of years and many evidences have shown that the technical analysis can help stock investor to make right judgment on stock market [2,3,19,21,22].

In academic research, many researchers also applied this approach in their forecasting models. For example, the study of charting patterns analysis for investment decision [19], investigating price charting patterns with kernel regression for the identification of ten patterns [16–18], and implementing a variation of the bull-flag stock chart with a template matching technique based on pattern recognition [31,33]. The stock chart patterns of these studies used for pattern recognition are established in fixed patterns given by experts and researchers, which are not very similar to the actual fluctuations in historical stock data, due to possible chart patterns contained historical stock data are not considered to forecasting.

Additionally, many advanced models and algorithms were proposed and these have achieved considerable results in forecasting accuracy. According to theories of model building, stock index forecasting models can be summarized in two categories [28]: (1) linear models based on statistical theories, such as General Autoregressive Conditional Heteroskedasticity (GARCH) and Stochastic Volatility model [12]; and (2) non-linear models based on artificial intelligence, such as the fuzzy time-series [5,29,30,34,35], Rough sets theory [5,14,23–26,30] artificial neural network [4,36,37], the support vector machine [32,38,39], and the particle swarm optimization [20,36,40].

After reviewing the related research, we argued that two issues are worth to be discussed further. Firstly, most researchers employed forecasting error as performance indicator, such as RMSE and MAPE, to evaluate their models [1,5,6, 28–30]. Although the daily forecasts are very useful for professional investors who implement intraday trading, we think that forecasting a bullish turning point is a more interesting issue than the future stock index because an accurate forecast for bullish stock pattern will bring a huge amount of stock return. Secondly, most stock pattern recognition models employed one matching template with many fixed weights assigned by researcher and it is not objective approach [8]; [16–18,31].

In this paper, we propose three new approaches to refine past methods: (1) a new definition for a bullish turning point with the probability of occurrence; (2) a new weighted method, based on the occurrence of observations, to produce dynamic weights for matching template; and (3) a new hybrid model based on PIP bull-flag pattern matching [11] and bull-flag template [18,31] for recognizing stock pattern.

To evaluate the profitability of the proposed model, we will give several trading criteria in experiments such as thresholds, stock holding period and stopping loss. In model verification, the NASDAQ composite index and Taiwan stock market weighted index (TAIEX) are taken as experimental datasets, and four advanced algorithms as benchmarks.

2. Related works

This section gives a brief review of methodology for two stock pattern identification methods: perceptually important point (PIP) identification matching and template matching.

2.1. Perceptually important point (PIP) identification matching

Reducing the dimension (i.e. the number of data point) by preserving the salient points is a promising method for time series representation [10]. These points are called as "*perceptually important points* (*PIP*)". The PIP identification process is introduced by Chung et al. [7] and used for matching of technical (analysis) patterns in financial applications.

In the PIP identification process, the vertical distance measure method is an automatic algorithm to recognize a specific pattern [11]. Three parameters are defined in the method: *fitting-window*, *holding-period*, and *distance-threshold*. The *fitting-window* is a window size of trading day for a stock pattern within a specific period of trading day such as 20-day, 40-day and 60-day. The *holding-period* is a certain period of trading day for investor to hold a stock. The *distance-threshold* is a specific price distance, *d*, for the method to define a stock pattern.

Fig. 1 illustrates the vertical distance measure method how to produce a price distance, *d*, for the "*head-and-shoulder*" pattern. In Fig. 1, three stock data, p_1, p_2 and p_3, are defined as three graphic points of two-dimension coordinates(x represents "a series of trading days" and y represents "stock price") contained in a specific period of *fitting-window*, where p_1