

## A NOVEL mechanism for forecasting based on pattern recognition

Hong Zhou<sup>1, a</sup>, Yuxiang Jin<sup>2</sup>, Yu He<sup>3, 4</sup>

<sup>1</sup> School of Finance, Renmin University of China, No. 59 Zhongguancun Street, Haidian District, Beijing, 100872, China

<sup>2</sup> Galaxy Derivatives Financial Service Co.Ltd, Beijing, 100045, China

<sup>3</sup> School of Finance, Zhejiang University of Finance and Economics, Hangzhou, 310018, China

<sup>4</sup> China Academy of Financial Research, Zhejiang University of Finance and Economics, Hangzhou, 310018, China

<sup>a</sup>email: zhouhong@ruc.edu.cn, sophie2000@263.net

**Keywords:** Pattern recognition; Trend forecasting; Stock market

**Abstract.** Trend prediction is an attractive topic, because determining the buy and sell points is one of the most important issues for investors in the stock market. The main purpose of this paper is to find an effective mechanism to identify the buy and sell signals. Based on the principle of pattern recognition, we establish a novel mechanism to recognize up or down patterns and the pinpoint for buying or selling. Using all stocks listed on the US NASDAQ, which includes 4460 companies for the period Jan. 1st 2010 to Dec. 31st 2010, our empirical study examines the predictive power of our trading strategies. The results indicate that our new strategies are profitable and that the accuracy rate of our predictions is above 67%. The main contribution of our study is that we find the trend to be more important than the isolated price. We first propose the novel pattern of “uptrend” or “downtrend”, while a continuous unidirectional trend may forecast the uptrend or downtrend in the future, two continuous unidirectional trends could verify the uptrend or down trend. That is, when a stock is on the rise, there must appear continuous up lines on the chart, and if the stock is falling, two down lines appear. We named this strategy as the “Two-line Strategy”

### Introduction

Stock market prediction is an attractive aspect of financial forecasting. Determining the buy and sell points is one of the most important issues for investors in the stock market. In practice, as the conventional technical analysis methods, such as the Dow Theory, the Elliott wave theory and the Gann are not very intuitionistic; it makes sense more in theory analysis and in practice it is difficult to use. The main purpose of this paper is to find an effective and intuitionistic mechanism to identify the buy or sell signal.

With the development of modern technology, a variety of modern technical analysis methods have emerged, such as time series analysis, the gray prediction model, the neural network model, and so on.

Generally, no matter what the conventional theories and modern methods, the close price is used as the most important parameter to predict the market. However, because of its delayed and hysteretic nature, the close price cannot timely reflect the reversal of the trend in the prices of securities. For example, as shown in Figure 1, the close price is slightly higher than the former price of the previous several days. The fact is that the trend has begun to reverse and the stock price has fallen. The hysteretic nature of the close price results in inaccuracies in the aforementioned indexes.

Further, as shown in Figure 1, the close prices of the first four continuous K lines are slightly increasing, while from the third K line, the highest price is gradually decreasing, until coming into the fall zone. These results indicate that the close price cannot timely reflect the trend of the stock market. Seeking to seize the turn point in time, we built a novel model to reflect the change of the trend in the stock market, and we argue that the change of the trend is more important than the changes in the close price. Accordingly, accurately defining the trend is the aim of this paper.

From the study, we conclude that the trend is more important than the isolate price and that continuous unidirectional lines may forecast an uptrend or downtrend. Studies have also found that the emergence of two continuous unidirectional lines is an important indicator of a short-term turning point. That is, when a stock is on the rise, continuous up lines will appear on the chart, while if the stock is on the decline, two down lines will appear on the chart.

Based on the principle of pattern recognition, we establish a novel mechanism to recognize upward and downward patterns and to identify the pinpoints for buying and selling. Using all stocks listed on the US NASDAQ, which includes 4460 companies for the period Jan. 1st 2010 to Dec. 31st 2010, our empirical study examines the predictive power of our trading strategies and concludes that our new pattern is profitable. Furthermore, conclusions have found that the accuracy rate of our prediction exceeds 67%.

The paper is organized as follows. We present the related literature in Section 2, and the model is described in Section 3. Section 4 discusses the empirical findings, and Section 5 presents the results and concludes the paper.

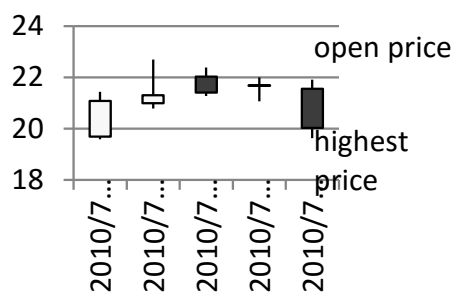


Fig.1 The simulated trend of stock price

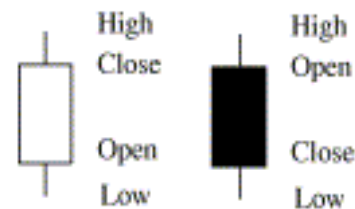


Fig. 2. Open, high, low and close prices displayed as candles.

## Methodology

Technical analysis as a reflection on the idea that prices move in trends is determined by the changing attitudes of investors toward a variety of economic, monetary, political, and psychological forces. Thus, technical analysis includes loads of forecasting techniques, such as chart analyses, cycle analyses and computerized technical trading systems. Because the processing power of computers was not available for statistical analysis in early times, early technical analyses were almost exclusively based on the analyses of charts. In sharp contrast to the views of many practitioners, some academics expressed skepticism about technical analyses even as more and more academics verified the profitability of such analyses. L. P. Lukac, et al. [1], A. W. Lo, et al. [2], E. R. Dawson and J. M. Steeley [3], and C.-H. Park and S. H. Irwin [4] are representative of these academics who reviewed the evidence on the profitability of technical analyses.

However, as more technical tools and theories have been developed and enhanced in recent decades, they have made an increasing impact on computer-assisted techniques that apply specially designed computer software, algorithms, and mathematical models, some recent studies have attempted to test visual chart patterns using pattern recognition algorithms. These modern methods include the model bootstrap and reality check methods, the random walk hypothesis method, the Kalman filter method, the grey theory, the fuzzification method, genetic programming, the artificial neural network method, etc.

Meanwhile, some academics combined candlestick techniques with modern technology to predict the stock market. For example, K.-i. Kainijo and T. Tanigawa [5] and M. Jasemi, et al. [6] applied recurrent neural networks to recognize stock price patterns and to develop a new method for evaluating the networks. Y.-M. Shiu and T.-H. Lu [7], T.-H. Lu, et al. [8], using the Taiwan 50 index component stocks for the period Jan. 2, 2002 to Dec. 31, 2009, investigated the predictive power of the candlestick two-day patterns and attempted to determine the key factors behind improving performance. They proposed the following three confirmation factors: the opening on the day after

a reversal pattern, the changes in real bodies between the two days, and the changes in volume. C.-H. L. Lee, et al. [9], C.-H. L. Lee, et al. [10] proposed an approach to extract fuzzy candlestick patterns from a financial time series and to select a set of patterns for investment decision making. T.-H. Lu, et al. [8] investigated the profitability of two-day candlestick patterns by buying based on bullish (bearish) patterns and holding until bearish (bullish) patterns occurred. They found that three bullish reversal patterns are profitable in the Taiwan stock market. A. Zapranis and P. E. Tsinaslanidis [11] also identified a rigorous rule-based mechanism for identifying the rounding bottoms (also known as saucers) pattern and resistant levels. Recently some academics used AI to help price finding, such as G. Caginalp et al. [12], J. H. Fock et al. [13], B. R. Marshall et al. [14], M. J. Horton [15], P. Roy, S. Sharma et al. [16].

We think that stock market movement trends are composed rising trends and falling trends. A large rising trend is composed of several small risings trends, while a falling trend is comprised of several small falling trends. Based on pattern recognition theory and graphic recognition theory, we must first identify a correct sample pattern, and because upward trends have certain common properties, we must build a correct model to present it. The same holds for downward trends. Furthermore, as investors, they should not only take advantage of the trends, but more importantly, they should identify the timing of the trend turning point, which is the focus of this paper.

As the change in the stock price becomes more intuitionistic in the candlestick chart, we use this chart define the basic pattern, as shown in Fig.2. Over a regular time interval (such as one minute, every fifteen minutes, one hour, one day, etc.), we define the beginning price as the open price and the ending price as the close price. If the close price is higher than the open price, we define the status as an uptrend. The uptrend is depicted by the white bar. In contrast, if the close price is lower than the open price, we define the status as a downtrend, which is depicted by the black bar. When the close price is higher (lower) than the open price, the candle “bar” is white (black).

Based on the observation and study of the movement track of the stock market, we find that the emergence of an uptrend and a downtrend has strong potential significance for accurate prediction. Thus, the value of this uptrend or downtrend is more important than the close price. The continuous small uptrends can constitute a significant uptrend, while the continuous small downtrends can constitute a significant downtrend. In particular, the continuous emergence of two white bars or two black bars often results in the turning point of the trend. Generally, because the close price is not the highest (lowest) price over a certain period of time, we use the highest price and lowest price in our model to determine trend movements.

Studies have shown that the emergence of the two continuous white bars (white lines) or black bars (black lines) can be used as an important indicator of a short-term turning point. That is, the emergence of two continuous reverse lines can be perceived as the turning point.

## Modeling

Given the event, “turn of trend”, denoted by A, and sample space, denoted by  $\Omega$ , are the causes of the K-line graph model of event A.

$$\Omega = [B_1, B_2, \dots, B_n] T$$

Here,  $B_i$  is composed by the different K-line graphs. That is,  $B_i \cap B_j = \Phi$ ,  $i \neq j$ ,  $\bigcup_{i=1}^n B_i = \Omega$  and we assume  $P(B_i) > 0$ .

According to the Bayesian formula, the possibility of event A caused by a certain K-line is:

$$p(B_i | A) = \frac{P(B_i)P(A|B_i)}{\sum_{i=1}^n P(B_i)P(A|B_i)}$$

That is, it is assumed that the emergence of a white line after continuous black lines or the emergence of a black line after continuous white lines will result in the reversal in the level of the date line. As the possibility of this event is expressed as  $P(A|B_1)$ , the possibility of the reversal of the trend is:

$$P(A) = P(A|B_1) * P(B_1) + (1 - P(A|B_1)) * P(B_2)$$

$$\text{and } P(B_1|A) = P(B_1) * P(A|B_1) / (P(A|B_1) * P(B_1) + (1 - P(A|B_1)) * P(B_2)).$$

Assumption:

- 1) Event A denotes the emergence of the reversal;
- 2) The graphic models of the emergence of a black line after continuous white lines or the emergence of a white line after continuous black lines are denoted as  $B_1, B_2, \dots, B_n$ ;
- 3) The models of the emergence of two continuous black lines after continuous white lines or the emergence of two continuous white lines after continuous black lines are denoted as  $B_1$ .

Accordingly, the purpose of this paper is to quickly and accurately capture the market turning point and predict the reversal of the white lines after continuous black lines and vice versa.

## Empirical analysis and test statistics

For general use and to verify that this model fits all time periods regardless of bull market or bear market, we select, from the US securities market, daily prices between Jan. 1<sup>st</sup> 2010 to Dec. 31<sup>th</sup> 2010 as study samples for this paper. Accordingly, there are 4460 listed companies in the *NASDAQ* during our sample period.

Because the daily prices are an important tool when making investment decisions, in the empirical study, daily trading is selected as a critical aspect in the formation of the whole trend. For other trading ranges, such as the 5-minute-trading-interval, 15-minutes-trading-interval, 5 day-trading-interval, one-week trading-interval, etc., the proposed model is equally applicable.

To specify the significance of two continuous white lines and two continuous black lines, two types of strategies are refined.

The first strategy is termed Strategy I. If there is a stock demonstrating characteristics of an uptrend, it must show a series of white bars. However, if a black bar appears, and if the next bar is also black, it indicates that the stock may well be entering the fall zone. The first bar is a white bar, suggesting that the stock is on an uptrend, but the second bar emerges and begins to change into a black bar. Such a change suggests the possibility of a trend reversal, and when the third bar is also black, it signals the selling of the stock as the stock would likely be down in the next time interval.

The second strategy, termed Strategy II, is an extension of Strategy I. If there is a stock demonstrating an upward trend, it must show a series of white bars. However, if a black bar appears and is then followed by another black bar, it suggests that the stock is entering the fall zone. Accordingly, there may be a signal to predict the next bar by comparing the highest prices. If the highest price of the second bar is lower than the highest price of the first bar, it would indicate an upcoming fall. The first bar is white, thus suggesting an uptrend. However, when the second bar emerges, it begins to change into a black bar, suggesting the possibility of a trend reversal. Because the highest value of the second bar is less than the highest value of the first bar, it signals the selling of the stock as the downtrend appears evident. In other words, the signal is given before the third bar appears. If the third bar is also a black bar, thus supporting our interpretation of the signal, our prediction, and our suggestion to sell as the stock would fall in the next time interval.

The data selection in the paper is from the US *NASDAQ*, which includes 4460 listed companies for the period Jan. 1<sup>st</sup> 2010 to Dec. 31<sup>st</sup> 2010.

## Strategy I applied in *NASDAQ* Stock Market

We calculate the accuracy rates for each listed company in the *NASDAQ* according to Strategy I.

Table 1 shows the descriptive statistics for Strategy I. To examine the effectiveness of this strategy, we use the T-test. Table 2 shows the T-test result.

Assume:  $\mu$  denotes the accuracy rate

$H_0: \mu \leq 65$   $H_1: \mu > 65$

As  $P=0 < \alpha=0.05$ , we deny the assumption  $H_0$ .

From the table, it shows that the accuracy rate is significantly greater than 65%. From the study, the results indicate that the signal for buy (sell) has no relationship with the number of continuous

days before the buy (sell) day. Meanwhile, the results also indicate that the highest price (lowest price) has greater significance than the close price. The above results verify that Strategy I is effective and profitable.

Table.1 Descriptive statistics of Strategy I<sup>1</sup>

|                              | N    | Minimum | Maximum | Average  | Standard deviation | Skewness | Kurtosis |
|------------------------------|------|---------|---------|----------|--------------------|----------|----------|
| black to white accuracy rate | 4314 | .00%    | 100.00% | 68.0570% | 10.64762%          | -.523    | 1.641    |
| white to black accuracy rate | 4314 | .00%    | 100.00% | 65.6248% | 11.02255%          | -.470    | 1.488    |

Table.2 T-test of strategy I

| Value=65%                    | T      | Df   | Sig.(two-sided) | Mean difference | 95% Confidence interval |             |
|------------------------------|--------|------|-----------------|-----------------|-------------------------|-------------|
|                              |        |      |                 |                 | Lower limit             | Upper limit |
| black to white accuracy rate | 18.858 | 4313 | .000            | 3.05703%        | 2.7392%                 | 3.3748%     |
| white to black accuracy rate | 3.723  | 4313 | .000            | .62484%         | .2958%                  | .9538%      |

## Strategy II applied to the NASDAQ

In the same way, programming by VB, we calculate the accuracy rates for each listed company in the *NASDAQ* according to the second strategy.

Table3 shows the descriptive statistics for Strategy II. To further examine the effectiveness of this strategy, we use the Kolmogorov-Smirnov test and the t-test.

Table.4 shows the Kolmogorov-Smirnov test of Strategy II. From the K-S test, both of the P values are less than 0.05, thus proving that the accuracy rates are normally distributed. As we are more concerned about whether the accuracy rates can represent the population condition, we use the t-test to verify the mean values of the accuracy rate groups, to determine whether they have significant differences and whether they are significantly greater than an absolute high value (67%), thus proving that the proposed strategy can be applied to different companies.

Table 5 shows the t-test of Strategy II.

Assume:

$\mu$  denotes the accuracy rate

$H_0: \mu \leq 67$   $H_1: \mu > 67$

As  $P=0 < \alpha=0.05$ , we should deny the assumption  $H_0$ .

Therefore, we can conclude that the accuracy rate is significantly greater than 67%. It also shows that the signal for buy (sell) has no relationship to the number of continuous days before the day. Meanwhile, the results indicate that the highest price (lowest price) has more value than the close price. The above results verify Strategy II is effective and profitable.

Table.3 Descriptive statistics for Strategy II<sup>2</sup>

|                              | N    | Minimum | Maximum | Average  | Standard deviation | Skewness | Kurtosis |
|------------------------------|------|---------|---------|----------|--------------------|----------|----------|
| Black to white accuracy rate | 4384 | 16.67%  | 100.00% | 71.1618% | 7.74446%           | -0.253   | 2.009    |
| white to black accuracy rate | 4384 | 16.67%  | 100.00% | 67.4349% | 7.96256%           | -0.247   | 2.074    |

Table.4 Kolmogorov-Smirnov test of Strategy II

|                      | Green to Red accuracy rate | Red to Green accuracy rate |
|----------------------|----------------------------|----------------------------|
| Kolmogorov-Smirnov Z | 2.941                      | 1.770                      |
| P                    | .000                       | .004                       |

Table.5 T-test of Strategy II

| Value=65%                    | T      | Df   | Sig.(two-sided) | Mean difference | 95% Confidence interval |             |
|------------------------------|--------|------|-----------------|-----------------|-------------------------|-------------|
|                              |        |      |                 |                 | Lower limit             | Upper limit |
| black to white accuracy rate | 35.582 | 4383 | .000            | 4.16185%        | 3.9325%                 | 4.3912%     |
| white to black accuracy rate | 3.617  | 4383 | .000            | .43492%         | .1991%                  | .6707%      |

According to the above two prediction strategies, we establish two corresponding trade strategies to execute imitation trading. And we compared with RSI, MACD and KDJ strategies, which are popular indicators used by investors. Yield results are list in Table 6 and 7.

<sup>1</sup>For some stock, there is little adverse point (less than 3 times) in one year, these data is meaningless, so we remove these incorrect data, then there is a little difference with the Table.1.

<sup>2</sup>For some stock, there is little adverse point (less than 4 times) in one year, these data is meaningless, so we remove these incorrect data, then there is a little difference with the Table.4.

Table.6 Yield results compared in Strategy I

|           | total times | average yield | yield per trade |
|-----------|-------------|---------------|-----------------|
| Strategy1 | 70471       | 115.65%       | 7.33%           |
| Strategy2 | 73133       | 185.74%       | 11.34%          |
| RSI       | 20971       | 17.50%        | 3.68%           |
| MACD      | 4801        | -2.68%        | -2.50%          |
| KDJ       | 504         | 1.54%         | 13.62%          |

Table.7 Yield results compared in Strategy II

|           | total times | average yield | yield per trade |
|-----------|-------------|---------------|-----------------|
| Strategy1 | 70471       | 9.00%         | 0.57%           |
| Strategy2 | 73133       | 15.18%        | 0.92%           |
| RSI       | 20971       | 17.50%        | 3.68%           |
| MACD      | 4801        | -2.68%        | -2.50%          |
| KDJ       | 504         | 1.54%         | 13.62%          |

From the result the average return based on either Strategy I or Strategy II yields a good profit margin. Compared with other indicators, such as RSI, MACD, KDJ, our strategies are effective.

Because we do not make a distinction between the bull market and the bear market on long-term trends in the market, our strategies can fit all market conditions, including a bull or a bear market.

## Conclusion

The main purpose of this paper is to find an effective and intuitionistic mechanism to identify the buy or sell signal. Based on the charting analysis method and the principle of pattern recognition, we establish a novel mechanism to recognize upward and downward patterns and the pinpoints for buying and selling.

The studies have found that the emergence of two red lines after continuous green lines indicates a reverse in the stock trend and vice versa. That is, the emergence of two opposite continuous K-lines can indicate a trend reversal, a strategy known as the "Two-line Strategy".

our empirical study examines the predictive power of our trading strategies and finds that our new strategies are, indeed, profitable. The results also indicate that our accuracy rate for predicting stock price changes exceeds 67%, thus suggesting that when compared with other indicators, such as RSI, MACD, KDJ, our strategies are effective and profitable.

The results of the empirical study further indicate that two continuous red lines usually imply an upward trend, while two continuous green lines usually imply a downward trend. In an upward trend, the low price forecast is higher than the high and the close prices, while in a downward trend, the high price forecast is higher than the low and the close prices.

The main contribution of our study is that we find that the trend is more important than the price, and we are the first to propose the novel pattern of "uptrend" and "downtrend", that is, the "Two-line Strategy". A continuous unidirectional trend may forecast a future uptrend or downtrend, especially in that two continuous unidirectional trends could confirm either an ascending or a descending trend. In other words, when a stock is increasing, continuous up lines appear in the chart, while when the stock is declining, two down lines appear in the chart.

## References

- [1] L. P. Lukac, B. W. Brorsen, and S. H. Irwin, "A Test of Futures Market Disequilibrium Using Twelve Different Technical Trading Systems," *Applied Economics*, vol. 20, pp. 623-639, 1988.
- [2] A. W. Lo, H. Mamaysky, and J. Wang, "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation," *The Journal of Finance*, vol. 55, pp. 1705-1770, 2000.
- [3] E. R. Dawson and J. M. Steeley, "On the Existence of Visual Technical Patterns in the UK Stock Market," *Journal of Business Finance & Accounting*, vol. 30, pp. 263-293, 2003.
- [4] C.-H. Park and S. H. Irwin, "What do we know about the profitability of technical analysis?," *Journal of Economic Surveys*, vol. 21, pp. 786-826, 2007.

- [5] K.-i. Kainijo and T. Tanigawa, "Stock Price Pattern Recognition - A Recurrent Neural Network Approach," pp. 215-221, 1990.
- [6] M. Jasemi, A. M. Kimiagari, and A. Memariani, "A modern neural network model to do stock market timing on the basis of the ancient investment technique of Japanese Candlestick," *Expert Systems with Applications*, vol. 38, pp. 3884-3890, Apr. 2011.
- [7] Y.-M. Shiu and T.-H. Lu, "Pinpoint and synergistic trading strategies of candlesticks," *International Journal of Economics and Finance*, vol. 3, pp. 234-244, Feb. 2011.
- [8] T.-H. Lu, Y.-M. Shiu, and T.-C. Liu, "Profitable candlestick trading strategies—The evidence from a new perspective," *Review of Financial Economics*, vol. 21, pp. 63-68, 2012
- [9] C.-H. L. Lee, Y.-C. Liaw, and L. Hsu, "Investment decision making by using fuzzy candlestick pattern and genetic algorithm," presented at the Fuzzy Systems (FUZZ), IEEE International Conference on Date 27-30 June 2011.
- [10] C.-H. L. Lee, A. Liu, and W.-S. Chen, "Pattern discovery of fuzzy time series for financial prediction," presented at the Knowledge and Data Engineering, *IEEE Transactions on Vol.18*, Issue: 5, 2006, pp. 613-625
- [11] A. Zapranis and P. E. Tsinaslanidis, "A novel, rule-based technical pattern identification mechanism: Identifying and evaluating saucers and resistant levels in the US stock market," *Expert Systems with Applications*, vol. 39, pp. 6301-6308, 2012.
- [12] G. Caginalp and H. Laurent, "The predictive power of price patterns," *Applied Mathematical Finance*, vol. 5, pp. 181-206, 1998.
- [13] J. H. Fock, C. Klein, and B. Zwergel, "Performance of Candlestick Analysis on Intraday Futures Data," *Journal of Derivatives*, vol. 13, pp. 28-40, 2005.
- [14] B. R. Marshall, M. R. Young, and L. C. Rose, "Candlestick technical trading strategies: Can they create value for investors?," *Journal of Banking & Finance*, vol. 30, pp. 2303-2323, Aug. 2006.
- [15] M. J. Horton, "Stars, crows, and doji: The use of candlesticks in stock selection," *The Quarterly Review of Economics and Finance*, vol. 49, pp. 283-294, May 2009.
- [16] P. Roy, S. Sharma, and M. K. Kowar, "Fuzzy Candlestick Approach to Trade S&P CNX NIFTY 50 Index using Engulfing Patterns," *International Journal of Hybrid Information Technology*, vol. 5, pp. 57-66, July 2012.