Generating Trading Strategies Based on Candlestick Chart Pattern Characteristics

Siriporn Thammakesorn and Ohm Sornil
School of Applied Statistics, National Institute of Development Administration, Bangkok, Thailand
E-mail: siriporn.tha@stu.nida.ac.th

Abstract. Candlestick chart patterns are widely used in stock trading decisions. Patterns of candlestick series are found to provide hints for the price of the next one. This research proposes a technique to generate stock trading strategies employing features which are shown to effectively recognize patterns in candlestick charts. The features are combined into a tree-like trading strategy using the Chi-square Automatic Interaction Detector algorithm. The technique is evaluated using actual stocks from Stock Exchange of Thailand. The results show that the generated strategies are more profitable than other popular trading techniques, such as moving average convergence divergence, exponential moving average, relative strength index, stochastic oscillator and average directional index.

1. Introduction
An investor expects a high return from investing in stocks. Stock prices are affected by many factors, such as economic conditions, political policies, natural disasters, etc. Fundamental analysis and technical analysis are two main approaches which form the basis of most traders’ decisions. Fundamental analysis considers economic, political, and detailed studies of companies’ financial positions while technical analysis employs indicators based on stock prices and volumes. Candlestick chart [1] is a popular method to display stock price information which consists of open price, high price, low price and close price. Patterns in candlestick charts are found to provide hints for future prices. They were invented in Japan with a long history for trading grains. Nowadays, they are applied to trading other instruments: stocks, futures and currencies. Lu and Shiu [2] study the profitability of candlestick patterns on Taiwan 50 Index component stocks from 2002 to 2009. They find that certain bullish candlestick patterns consistently outperform others. Moreover, they notice that buying signals are generally more effective than selling signals. Goo et al. [3] analyze daily data of 25 component stocks in the Taiwan Top 50 Tracker Fund and Taiwan Mid-Cap 100 Tracker Fund from 1997 to 2006. They find a strong evidence that certain candlestick trading strategies are profitable. In addition, they notice that different candlestick patterns require different holding periods to be profitable. There are more than 60 types of candlestick patterns which make it difficult for human to apply them effectively in investments. Some researchers study the use of candlestick charts and artificial intelligence in stock trading. Izumi, Yamauchi, Mabu, Hirase and Hu [4] create trading rules by combining genetic network programming (GNP) with candlestick charts. It is found that the average profit margin derived from GNP is higher than others. It can be seen that adopting a candlestick chart to create a trading rule can help increase return on investment. Chalothorn Chootong [5] create trading signals by combining a neural network (NN) with indicators, price chart patterns, and candlestick chart patterns. It can be seen
that candlestick chart patterns can help investors make better decisions. Leslie C.O. Tiong, David C.L. Ngo, and Yunli Lee [6] used candlestick patterns to study two AI techniques: artificial neural network (ANN) and support vector machine (SVM) in learning patterns and behaviours of stock prices for future price prediction. It was found that the best technique can predict the stock price correctly with 60 percent accuracy. Our previous research [7] discovers a set of features that can recognize candlestick chart patterns effectively. In this research, a trading strategy is generated from those features using the Chi-square Automatic Interaction Detector (CHAID) algorithm. The goal is to create trading strategies that provide high returns on investment. Experiments are conducted on 10 stocks from Stock Exchange of Thailand over ten years period, and the results are compared with popular trading strategies in details.

In the rest of the paper, the proposed technique is presented including candlestick chart pattern characteristics; results of the experimental evaluations are shown and discussed; and finally conclusions are drawn.

2. Proposed Technique

Our study is shown in Figure 1. Training data is used by the Chi-square Automatic Interaction Detector (CHAID) algorithm to build a trading strategy in form of a classification tree using features which can effectively recognize candlestick chart patterns. The model is then tested with actual price data of unseen periods. This section first introduces candlestick charts, the features used in the proposed method and generation of a trading strategy.

2.1. Candlestick Chart Characteristics

Candlestick charts [8, 9] originated from Japan. It has a history of more than 200 years for analyzing rice and grain prices. After several decades, the chart is adopted into trading securities and becomes a very popular technical analysis technique. A candlestick (see Figure 2.) shows relationships between open price, high price, low price and close price.

Patterns of candlesticks are found to provide hint on future prices. However, there are more than 60 patterns making it difficult for an investor to use them effectively. In our previous work [8], a set of 26 features for automatic recognition of candlestick chart patterns is proposed and evaluated. The features
are based on colours, sizes and relative positions of candlesticks. We can then make use of those features to construct a trading strategy for predicting future prices. The features consist of:

- Prior Trend
- Colour of the 1st candlestick
- Colour of the 2nd candlestick
- Colour of the 3rd candlestick
- Size of the 1st candlestick normalized by the size of the largest candlestick
- Size of the 2nd candlestick normalized by the size of the largest candlestick
- Size of the 3rd candlestick normalized by the size of the largest candlestick
- The ratio between the sizes of the 2nd and 1st candlesticks
- The ratio between the sizes of the 3rd and 1st candlesticks
- The ratio between the sizes of the 3rd and 2nd candlesticks
- Gap between the 1st and 2nd candlesticks
- Gap between the 1st and 3rd candlesticks
- Gap between the 2nd and 3rd candlesticks
- Real body gap between the 1st and 2nd candlesticks
- Real body gap between the 1st and 3rd candlesticks
- Real body gap between the 2nd and 3rd candlesticks
- Slope between the low price between 1st and 2nd candlesticks
- Slope between the high price between 1st and 2nd candlesticks
- Slope between the low price between 1st and 3rd candlesticks
- Slope between the high price between 1st and 3rd candlesticks
- The real body size of the 1st candlestick normalized by the largest real body size
- The real body size of the 2nd candlestick normalized by the largest real body size
- The real body size of the 3rd candlestick normalized by the largest real body size
- The real body size of the 1st candlestick normalized by its total size
- The real body size of the 2nd candlestick normalized by its total size
- The real body size of the 3rd candlestick normalized by its total size

2.2. Generating Trading Strategy
A trading strategy is a classification tree with 3 output values: Buy, Sell or Hold. The strategy is created by the Chi-square Automatic Interaction Detector (CHAID) algorithm [10]. CHAID uses Pearson’s Chi-square test of independence, which tests for an association between two categorical variables, to determine the p-value which is used to select the variables to split a node. The statistics for the chi-square test of independence is

\[ \chi^2 = \sum_{i,j} \left( \frac{O_{ij} - E_{ij}}{E_{ij}} \right)^2 \]

where \( i \) is 1, 2, ..., \( r \) and \( j \) is 1, 2, ..., \( c \), \( O_{ij} \) is the observed (actual) frequency value, and \( E_{ij} \) is the expected frequency value. CHAID consists of three steps, i.e., merging, splitting and stopping as follow:

- Merging
  This step will calculate a significant test of each categorical independent variable towards the dependent variables and merge most similar categories of independent variables by selecting the smallest significant or the biggest p-value of a category.

- Splitting
  Splitting step separates a node by considering the biggest significant or the smallest p-value of independent variables obtained from the merging process.

- Stopping
The stopping step checks for terminating the creation and separation of nodes, with the following requirements:

a) If the depth of the tree reaches the maximum depth limit, the data classification process stops.
b) If the size of a node is less than the minimum node size, then the node will stop splitting.

3. Experimental Evaluations
To evaluate the proposed method, data of 10 stocks from Stock Exchange of Thailand (SET) during 2008 to 2017 are used in the experiments. The data are divided into training data (from 2008 to 2015) and testing data (from 2016 to 2017). Amibroker software [11] is used for backtesting and measuring effectiveness of each strategy. Buying and selling orders will be issued at the opening price on the next day after the day yielding the signal. The proposed method is compared with 8 popular trading strategies which are:

- Buy and Hold
- EMA: If EMA (5) crossovers EMA (20) then buy. If EMA (20) crossovers EMA (5) then sell.
- MACD (1): If Fast (12) crossovers Slow (26) then buy. If Slow (26) crossovers Fast (12) then sell.
- MACD (2): If MACD (12,26) crossovers Signal (9) then buy. If Signal (9) crossovers MACD (12,26) then sell.
- RSI: If RSI (14) is less than 30 then buy. If RSI (14) is greater than 70 then sell.
- STO: If Fast (14) crossovers Slow (3) then buy. If Slow (3) crossovers Fast (14) then sell.
- STO: If Slow (14, 3) is greater than 20 then buy. If Slow (14, 3) is less than 80 then sell.
- ADX: If +DI crossovers -DI then buy. If -DI crossovers -DI then sell.

Table 1. Results of Backtesing

| Stock   | Trading Method | Net Profit (Percent) | Trading Number | Stock   | Trading Method | Net Profit (Percent) | Trading Number |
|---------|----------------|----------------------|----------------|---------|----------------|----------------------|----------------|
| ADVANC  | Propose Method | 36.04%               | 90             | AMARIN  | Propose Method | 8.92%                | 69             |
| Buy and Hold | 33.57%            | 1                    | EMA            | Buy and Hold | -16.36%        | 1                  |
| EMA     | 20.68%          | 12                   | MACD(1)        | EMA     | -14.15%        | 21                 |
| MACD(1) | -4.88%          | 7                    | MACD(2)        | MACD(1)  | -17.69%        | 8                  |
| MACD(2) | 28.14%          | 15                   | RSI            | MACD(2)  | -21.21%        | 22                 |
| RSI     | 25.78%          | 3                    | STO(1)         | RSI     | -8.38%         | 2                  |
| STO(1)  | 32.10%          | 56                   | STO(2)         | STO(1)  | 10.96%         | 56                 |
| STO(2)  | 15.66%          | 9                    | ADX            | STO(2)  | -16.53%        | 3                  |
| ADX     | 18.76%          | 12                   |                | ADX     | -4.73%         | 16                 |

| AOT     | Propose Method | 36.32%               | 98             | BBL     | Propose Method | 34.67%              | 1              |
|---------|----------------|----------------------|----------------|---------|----------------|----------------------|----------------|
| Buy and Hold | 98.25%            | 1                    | EMA            | Buy and Hold | 34.67%        | 1                  |
| EMA     | 57.73%          | 11                   | MACD(1)        | EMA     | 14.96%         | 17                 |
| MACD(1) | 57.02%          | 5                    | MACD(2)        | MACD(1)  | 25.57%         | 6                  |
| MACD(2) | 30.59%          | 19                   | RSI            | MACD(2)  | -6.54%         | 23                 |
| RSI     | 14.29%          | 1                    | STO(1)         | RSI     | 7.19%          | 1                  |
| STO(1)  | 21.18%          | 63                   | STO(2)         | STO(1)  | 6.78%          | 58                 |
| STO(2)  | 8.35%           | 6                    | ADX            | STO(2)  | 4.49%          | 7                  |
| ADX     | 53.48%          | 19                   |                | ADX     | 4.63%          | 26                 |

| BKI     | Propose Method | 8.40%                | 80             | CPF     | Propose Method | 5.09%              | 111             |
|---------|----------------|----------------------|----------------|---------|----------------|----------------------|----------------|
| Buy and Hold | 0.85%            | 1                    | EMA            | Buy and Hold | 36.36%        | 1                  |
| EMA     | -4.73%          | 22                   | MACD(1)        | EMA     | -0.44%         | 15                 |
| MACD(1) | -0.77%          | 11                   | MACD(2)        | MACD(1)  | 11.33%         | 8                  |
| MACD(2) | -6.22%          | 23                   | RSI            | MACD(2)  | -1.41%         | 19                 |
| RSI     | 2.59%           | 2                    | STO(1)         | RSI     | -4.50%         | 1                  |
| STO(1)  | -8.34%          | 67                   | STO(2)         | STO(1)  | 0.72%          | 59                 |
| STO(2)  | -4.18%          | 5                    |                | STO(2)  | -7.29%         | 5                  |
The results of testing with data of 10 stocks for 2 most recent years (2016 to 2017) are shown in Table 1. Net Profit is the net profit achieved after completing all trades in the period, and Max System Drawdown is the largest peak to valley percentage decline experienced while trading in the period. We can see that strategies generated from the proposed technique yield the highest returns or are ranked among the top strategies for stocks that are in an upward trend. Also the generated strategies suffer the lowest losses or are among the top strategies for stocks that are in a downward trend. For some stocks, there may be a very few strategies that yield higher returns than our strategies; but there is no common strategy that generally outperforms our technique. Overall, strategies generated from the proposed technique outperform 8 other comparative techniques which mean that our technique is applicable to trading actual stocks.

4. Conclusion
Candlestick chart is a popular method to display stock price information. Patterns of candlestick series are found to provide hints for the price of the next one. There are so many types of candlestick chart patterns which are difficult to be used effectively. In our work, twenty six features for recognizing candlestick chart patterns are identified. This paper presents a technique to generate effective trading strategies based on those features. A trading strategy is a classification tree generated by the Chi-square Automatic Interaction Detector (CHAID) algorithm. The technique is compared with eight popularly used strategies on 10 stocks from Stock Exchange of Thailand. The results show that the proposed technique generates trading strategies that outperform other strategies, and the proposed technique is applicable to real-world security trading.

5. References
[1] Nison S 2001 Japanese Candlestick Charting Techniques (2nd ed.) New Jersey (Pearson Education US).
[2] Lu T-H and Shiu Y-M 2012 Tests for two-day candlestick patterns in the emerging equity market of Taiwan Emerging Markets Finance & Trade 48 41-57.
[3] Goo Y-J, Chen D-H and Chang Y-W 2007 The application of Japanese candlestick trading strategies in Taiwan Investment Management and Financial Innovations 4(4) 49-79.
[4] Izumi Y, Yamaguchi T, Mabu S, Hirasawa K and Hu J 2006 Trading Rules on The Stock Markets Using Genetic Network Programming with Candlestick Chart 2006 IEEE Congress on Evolutionary Computer 43 317-322.
[5] Chootong C and Sornil O 2012 Trading Signal Generation Using A Combination of Chart
Patterns and Indicators *International Journal of Computer Science Issues* 9 202-209.

[6] Tiong L C O, Ngo D C L and Lee Y 2013 Stock Price Prediction Model using Candlestick Pattern Feature *International Journal Of Interactive Digital Media* 1(3) 58-64.

[7] Dhammathanapatchara T and Sornil O 2015 Automatic Candlestick pattern recognition. *International Journal of Applied Engineering Research* 10 1833-1836.

[8] Bigalow S W 2011 *Profitable Candlestick Trading: Pinpointing Market Opportunities to Maximize Profits (2nd ed.)* New Jersey (John Wiley & Sons, Inc.).

[9] Corbitt W A 2012 *Candlestick Charting Demystified* New York (McGraw Hill).

[10] Kass G V 1980 An Exploratory Technique for Investigating Large Quantities of Categorical Data *Applied Statistics* 29 119-127.

[11] Bandy H B 2012 *Introduction to AmiBroker (2nd ed.)* South Dakota (Blue Owl Press, Inc.).