BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Journal of Business Administration Research
http://ojs.bilpublishing.com/index.php/jbar

ARTICLE

An Approach to Find the Point of Buying Stock Based on Big Data

Yao Fu* **Congdian Cheng** **Lizhu Wang**

College of Mathematics and Systems Science, Shenyang Normal University, Shenyang, 110034, China

ARTICLE INFO

Article history
Received: 17 April 2020
Accepted: 23 April 2020
Published Online: 30 April 2020

Keywords:
Stock
Big data
Weighted moving average
Algorithm; profit

ABSTRACT

It is a research subject that has attracted a wide concern and study for a long time to find a suitable trading point of stock. From the views of big data and quantization technique, the paper tries to propose an approach, through the form of algorithm, based on big data analysis and linear weighted moving average curve, to find the point of buying stock, so that the trader would like to achieve the expected profit with a higher probability; and makes the digital experiment to further explain the approach and verify its performance. This work can promote the development of big data research and quantization technique, and can also provide a certain reference method for the trader making the technology analysis of the trade.

1. Introduction

Since China developed its stock market, the scale has continuously expanded, the number of quoted companies has steadily increased. As the institutional construction has gradually perfected, the participants have become more and more. By the development of many years, now the stock market has become an important pillar of China's socialist economy with Chinese characteristics, which extremely adapts to the development of China's economy, and greatly promotes its prosperity, development and stability. The trade of stock has also become a major investment channel for China's people. Stock market is a huge complex system. A lot of theoretical work is needed to do for how to effectively build the system. Therefore, the China's stock market has received extensive attention and research since it came into being in the 1990s. For example, D.W.Tian discussed whether the reform of the divisional of stock right has an impact on the effectiveness of China's stock market, which provided a reference for the formulating of relevant policies [1]. S.Wang et al. predicted the feasibility of changes of stock price by Markov chain [2]. X.K.Li used the Copula function to analyze the correlation of Shanghai stock market and Shenzhen stock market [3]. X.L.Ren et al. argued the spillover effect of fluctuations between international energy market and China's stock market, as well make an empirical analysis [4]. R.W.Lin et al., based on multi-dimensional interactive verification, tried to predict the trend of the stock market from multiple dimensions and tested the method by establishing a relevant model [5]. X.P.Teng et al. proposed a method of predicting stock price trend based on the deep multiple regression model [6].

The moving average is often used to show the average level of historical fluctuations of stock prices, thus, to a certain extent, reflects the future development trend of stock prices. Therefore moving averages are an import-

**Corresponding Author:*
Yao Fu,
College of Mathematics and Systems Science, Shenyang Normal University, Shenyang 110034, China;
Email: fuyao0404@163.com

DOI: https://doi.org/10.30564/jbar.v3i2.1811

ant tool for the technical analysis of stock trade, and it is an important aspect of the research area of stock market to explore buying point and selling point from the moving averages. B.J.Chen et al. analyzed the effectiveness of trading based on moving averages by designing two indicators, namely the prediction accuracy of the combination of buying point and selling point, and the cumulative rate of return [7]. B.B.Sun tested whether the moving average has the ability to obtain excess profits [8]. C.Y.Wang based on the theory of moving average analyzed the effectiveness of trend investment in the stock market [9].

Restricted by the technique of data processing, such as ingestion, storage and compute, the previous work that studies the buying point and selling point from the moving average has not been able to fully take advantage of the role of big data method and quantitative investment technology. Noting the situation, X.X.Xu et al., based on quantitative the technique, developed an algorithm to, through big data analysis and moving average curve, find the buy point of stock; and conduct a digital experiment to further explain the algorithm and verify its performance [10]. The main idea of the algorithm is to strenuously find a suitable scale δ so that, when the stock price is below the positive difference between the most recent daily average *MA* of a certain period and the scale δ, the buy of stock can obtain expected profits with a high probability. Therefore, we regard the region of stock price below the positive difference as a buying interval. In order to find the scale δ, we review the historical data of trades for a certain long period of time; and supposed that the stock is bought whenever its price is below (MA-δ),we test how many trading days can achieve the expected profit, and calculate the frequency of reaching the expected profit, which is believed as a success probability . If this probability reaches the expected level, the δ is regarded as the suitable scale. And otherwise, we change the value of δ and continue the preceding process to find it. The process completes until the suitable scale is found. The work of literature [10] makes full use of big data methods and quantification techniques, and to a certain extent, develops related work of previous research.

Due to the hysteresis of simple moving averages, the algorithm proposed by the literature [10] may be inefficient sometimes for it fails to make full use of the most recent information. On the other hand, due to no consideration of the loss with the failures in the process to find the suitable scale according to the expected success probability, the algorithm may be inefficient sometimes for the loss is too more. In view of this observation, the present work tries to develop the work of literature [10] by changing the simple moving average as the weighted moving average, and the finding suitable scale according to the expected success probability as the finding suitable scale according to the maximum total profit. We will, based on quantitative the technique, propose an algorithm to, through big data analysis and weighted moving average curve, find the buy point of stock; and conduct a digital experiment to further explain the algorithm and verify its performance.

2. Mathematical Programming

In order to clarify the problem that will be considered, we provide a preparation from the perspective of mathematical modeling in this section.

Let *S* be a stock. We denote P_i^l、P_i^h and P_i^c to represent the lowest price, highest price and closing price with trading day i respectively, and use $(i-j)$ to represent the *j*th trading day before trading day i . Here, $i-1$ refers to the first trading day before trading day $i,\cdots,$ $(i-j)$ refers to the *j*th trading day before trading day i . We call

$$\overline{P}(i)=\frac{1}{1+2+\cdots+l}\sum_{j=1}^{l}P_{i-j}^{c}(l-j+1)$$

as linear weighted average stock price with the previous l trading days of trading day i ; and call the smooth curve formed by the points $\{(i,\overline{P}(i))\}$ as linear weighted moving average curve, as weighted moving average for simpleness and clearness .

For stock S , assume that we buy one share in trading day i , and the period we hold shall not exceed f trading days. Use $O(i,f)$ to represent the profit in the future f trading days afterwards. We consider the problem how to make $O(i,f)\geq v$ with a high probability, and denote it as $\mathrm{S\text{-}SPMP}(f,v)$, abbreviated as $\mathrm{SPMP}(f,v)$. Here v is an appropriate positive real number, which indicates the expected profit we hope to obtain.

Next, we tentatively establish an approach to solve the problem and conduct a digital experiment to further explain the algorithm and verify its performance.

Remark 1: In order to be convenience, for buying stock mentioned in present work, the number of buy is understood as one share. In practice, for to use the approach we proposed, the number can be adjusted through multiplying by the corresponding coefficient.

3. Algorithm

In this section, we try to, through establishing the following algorithm SPMO, propose an approach to solve the problem $\mathrm{SPMP}(f,v)$.

Obviously all the stock investor hope that the stock bought can get more profit within a certain period of time. A most popular approach is to buy stocks at the price below a certain moving average. It is a difficult question to tackle with that at what position in the region of stock price below the moving average we should buy the stock for, on the one hand, that the difference between the moving average and the buying price is less will result in the profit of buy stock reducing; on the other hand, that the difference is more will result in the chance to buy stock reducing. So, it is the bottleneck for the investors buy stocks according the approach to determine the buying point.

Given a problem $\mathrm{SPMP}(f,v)$, let

$P_i^l, P_i^h, P_i^c, i=1,2,\cdots,l+L+f,$

be respectively the lowest price, highest price and closing price with trading day *i* for the latest $(l+L+f)$ historical data, and P_0 be the current price. We try to find an suitable $\delta>0$ basing on these data such that it is possible to guarantee the profit v with a high probability within the next f trading days that the investors buy stocks when $P_0<[\overline{P}(1)-\delta]$.

We use the intensive search method to accomplish this task. Take an appropriate small $\varepsilon>0$. Then, we replay the trades, in turn, on L trading days $i=f+1,f+2,\cdots,f+L$, for

$$\delta=\varepsilon,2\varepsilon,3\varepsilon,\cdots,<\max_{f+1\le i\le f+L}\{P_i^h\}$$

respectively. Assume P_i^l to be the current price when we replay the trade for trading day i. On the trading day i, we first consider whether the stock can be buy at a price lower than $[\overline{P}(1+i)-\delta]$, namely, $P_i^l<[\overline{P}(1+i)-\delta]$. If it is not possible, 0 is used as the profit of trading day i; otherwise, we further consider in two cases that the profit can reach v, namely $[\overline{P}(1+i)-\delta+v]<\max_{i+1\le j\le i+f}\{P_j^h\}$; and the profit can not reach v, namely, $[\overline{P}(1+i)-\delta+v]\ge\max_{i+1\le j\le i+f}\{P_j^h\}$. In the first case, v is regarded as the profit of this trading day. In the second case, $\{P_{i+f}^l-[\overline{P}(1+i)-\delta+v]\}$ is regarded as the profit of this trading day. For an δ determined, after the replaying trade completes for each of the L trading days, take the sum of the profits of each trading day as the total profit with δ for the L trading days, and indicated it as R_k, where $k=\frac{\delta}{\varepsilon}$. Finally, we find the k^* that satisfies $R_{k^*}=\max\{R_k\}$, and use $\delta^*=k^*\varepsilon$ as the δ we want to find.

From the states above, we can obtain the following observation. On the one hand, if the historical data used is less, namely l and L are too small, from the viewpoint of statistic, the found δ is too difficult to guarantee success of high probability, that is, when the investors buy stocks under the condition $P_0<[\overline{P}(1)-\delta]$, they do not easily obtain profit v within the next f trading days, so the amount of historical data used should be sufficiently large. On the other hand, it does not meet the actual needs that the speed of calculation is too slow, so the speed of calculation should be sufficiently fast. That is to say, in order to effectively achieve the above ideas, we must use the big data technology. For processing big data at high speed, it is impossible to do without computers. Next, according to the above ideas, we design an algorithm SPMO to solve $\mathrm{SPMP}(f,v)$ with high probability, which can be executed by a computer and can better illustrate our method too.

SPMO Algorithm

Question: $\mathrm{S\text{-}SPMP}(f,v)$.

Input:

$f,l,L,v,\varepsilon,P_i^l,P_i^h,P_i^c,\ i=1,2,\cdots,l+L+f;\ P_0.\ (f,l,L,v,\varepsilon>0)$

(Here, f is the number of the trading days within which the investor expects to profit v, l is the number of trading days for $\overline{P}(i)$, L is the number of training days; v is the expected profits; ε is a parameter variable representing a kind of operation gap; P_i^l is the lowest price on trading day i, P_i^h is the highest price, P_i^c is the closing price on trading day i; $l+L+f$ is the number of historical data used; P_0 is the current price.)

Output: Yes (The investor can buy stocks at the current price); No (The investor cannot buy stocks at the current price).

Process:

(1) Calculate $\overline{P}(i)\quad i=f+1,f+2,\cdots,f+L$.

(2) $k:=0;\ \delta:=0,\ \delta_0:=0$.

(3) $k:=k+1;\ \delta_k=\delta_{k-1}+\varepsilon,\ \delta=\delta_k$.

If $\max\{\overline{P}(i)-\delta-P_i^l;\ i=i+1,\cdots,L+f\}>0$, go to the step 4; otherwise, go to the step 7.

(4) Put $R_k:=0$, then go to the step 5.

(5) For $i=f+1,f+2,\cdots,f+L$, proceed to the following items.

① If $P_i^l\ge\overline{P}(i)-\delta$, return to the step 5.

  DOI: https://doi.org/10.30564/jbar.v3i2.1811

② If $P_i^l < \overline{P}(i) - \delta$, proceed the next item.

③ If $\overline{P}(i) - \delta + v < \max_{i+1 \le j \le i+f} \{P_j^h\}$, put $R_k := R_k + v$; otherwise, put $R_k := R_k + (P_{i+f}^c - \overline{P}(i) + \delta)$.

(6) After completing the loop, return to the step 3.

(7) Find k^* such that $R_{k^*} = \max\{R_k\}$. Then, make $\delta^* = \delta_{k^*}$.

(8) If $P_0 < \overline{P}(0) - \delta^*$, output Yes; otherwise, output No. Stop.

4. Experiment

We take TongHuaShun (300033) of Shanghai stock market as an instance. The experiment selects 270 trading days from September 6, 2018 to October 22, 2019; takes the daily highest price, lowest price and closing price for each trading day selected as the original experimental data; chooses $l = 60, L = 120, f = 30, v = 4.5, \varepsilon = 0.2$, and takes respectively each trading day of 30 trading days from July 23, 2019 to September 2, 2019 as the current trading day and the lowest price of the current trading day as the current price. Then, in turn, for each current trading day, input the related data into the MATLAB program of algorithm SPMO and conduct the experiment. See the following and Table 1 for details.

(1) Take July 23, 2019 as the current trading day, or say trading day 0. On $i = 0$ (July 23, 2019), 1(July 22, 2019), 2 (July 19, 2019), …, $(L + f + l)$ (September 06, 2018), input the related data into the MATLAB program of algorithm SPMO, then obtain the result: No, i.e. can't buy , the profit is 0.

(2) Take July 24, 2019 as the current trading day, or say trading day 0. On $i = 0$ (July 24, 2019), 1 (July 23, 2019), 2 (July 22, 2019),..., $(L + f + l)$ (September 07, 2018), input the related data into the MATLAB program of algorithm SPMO, then obtain the result: No, i.e. can't buy , the profit is 0.

......

(30) Take September 02, 2019 as the current trading day, or say trading day 0. On $i = 0$ (September 02, 2019), 1(August 30, 2019), 2 (August 29, 2019),..., $(L + f + l)$ (October 25, 2018), input the related data into the MATLAB program of algorithm SPMO, then obtain the result: No, i.e. can't buy , the profit is 0.

From table 1, it can be seen that among 30 trading days, 18 trading days can buy the stock TongHuaShun (300033), and can receive the profit v within the future30 trading days. The conclusion shows that the algorithm SPMO is effective and has certain feasibility.

Table 1. Summary of experimental results

Current date	Buy or not	Banefit v or not	Current date	buy or not	Banefit v or not
2019.7.23	No	----	2019.8.13	Yes	Yes
2019.7.24	No	----	2019.8.14	Yes	Yes
2019.7.25	No	----	2019.8.15	Yes	Yes
2019.7.26	No	----	2019.8.16	Yes	Yes
2019.7.29	Yes	Yes	2019.8.19	Yes	Yes
2019.7.30	No	----	2019.8.20	Yes	Yes
2019.7.31	Yes	Yes	2019.8.21	Yes	Yes
2019.8.1	Yes	Yes	2019.8.22	No	----
2019.8.2	Yes	Yes	2019.8.23	Yes	Yes
2019.8.5	Yes	Yes	2019.8.26	No	----
2019.8.6	Yes	Yes	2019.8.27	No	----
2019.8.7	Yes	Yes	2019.8.28	No	----
2019.8.8	Yes	Yes	2019.8.29	No	----
2019.8.9	Yes	Yes	2019.8.30	No	----
2019.8.12	Yes	Yes	2019.9.2	No	----

Remark 2: The experimental data above comes from the stock trading software of TongHuaShun. In the experiment, the lowest price of the current trading day is taken as the current price.

Remark 3: (1) The simulation experiment above is completed through the following two steps. Firstly, write the MATLAB program of algorithm SPMO. Secondly, for the related data selected, run the program to perform big data operations on a computer with model of Inspiron1427FT02 and CPU of Intel(R) Core(TM)2 Duo CPU T6500 @2.10GHz.

(2) The algorithm SPMO not only has many parametric variables, but also can quickly extract a large amount of data from the stock trading software, quickly carry out calculations, and timely answer the question whether investors can buy the related stock at the current price. So, it fully reflects the characteristics of big data technology.

5. Conclusion

Influenced by the boom of big data and quantitative investment, as well as the previous related research, especially the work of X.X.Xu et al.[10], the present work tries to based on quantitative the technique, established an approach to, through big data analysis and weighted moving average curve, find the buying point of stock, and then conducts an experiment to make a further illustration. The experiment made by us shows that the approach is effective. It is the issue worthy of further research to , from others ideas, develop the new algorithms for finding stock trading points with the big data and quantitative investment technique, as well as discuss the rate of return on the developed algorithm. We will continue to work hard. And meanwhile we extremely look forward to the present work can stimulate the interest of much friends so that more scholars can take delight in this research area together promoting the development of the research work of quantitative technology and trading technology!

Acknowledgements

This work is supported by the Natural Science Foundation of Liaoning province under Great 20170540821.

Reference

[1] D. W. Tian. An Empirical Study on the Effect of the Split Share Reform on the Effectiveness of China's Stock Market[J]. Finance & Economy, 2017(04): 144-147.

[2] S. Wang, J. Z. Qiao. Forecasting stock price by Markov Chain:Take Yili Group as the instance[J]. Inner Mongolia Statistics, 2017(04): 7-9.

[3] X. K. Li. Correlation analysis of Shanghai and Shenzhen stock markets based on Copula function[J]. Journal of Shaanxi University of Technology, 2017, 33(06): 75-81.

[4] X. L. Ren, Y. K. Xiao, W. Y. Sun. Volatility Spillover Effect Between International Energy Market and Chinese Stock Market[J]. Journal of Ocean University of China, 2018(06): 65-71.

[5] R. W. Lin, Y. Yang. Stock Market Trend Prediction: Based on Multi-Dimensional Cross-Validation Method[J]. Journal of Shanghai University of Internatinal Business and Economic, 2018, 25(06): 40-49.

[6] X. P. Teng, W. Zheng. Deep Multiple Regression Model for Stock Price Trend Foreca-sting[J]. Economic Research Guide, 2019(21): 71-74.

[7] B. J. Chen, W. J. Chen. Research on Moving Average Analysis and Its Trading Strategy[J]. Commercial Research, 2015(07): 73-79.

[8] B. B. Sun. Are Moving Average Rules Profitable? Evidence from Shanghai Composite Index[J]. The Journal of Quantitative & Technical Economics, 2005(02): 149-156.

[9] C. Y. Wang. Research on the effectiveness of trend investment in Shanghai share market and Shenzhen share market: A strategy based on the moving average[J]. Business Economy, 2019(09): 178-180.

[10] X. X. Xu, L. Z. Wang, C. D. Cheng. A Method of Finding Stock Buying Point Based onMoving Average Price and Big Data Analysis[J]. Statistics and Application, 2020, 9(1): 1-6.

DOI: https://doi.org/10.30564/jbar.v3i2.1811