A Novel Model of Stock Data Mining with M/G/1 Queue for Evaluation of Stock Crash

Qingzhen Xu[1,2*] and Feifei Zhang[1]

[1]*School of Computer Science, South China Normal University, Guangzhou 510631, China*
[2]*China Investment Securities, Shenzhen 518048, China*
E-mail:xqz1997@126.com

Abstract

Data mining is the process of searching the information from a large amount of data. In order to evaluate the stock crash this paper proposes general decrementing service M/G/1 queue system with multiple adaptive vacations to find information related to stock crash in data about Shanghai Composite Index. We use the probability generating function (P.G.F.) of stationary queue length and LST of waiting time, and their stochastic decomposition to calculate Existing money flow. Existing Money flow calculation model is improved based on the stationary queue length and LST of waiting time. We program to achieve the stock of existing money flow algorithm, and get the number of existing money flow. The improved algorithm can early warn the stock market crash. The empirical result shows that: There will be a rise in price before the Stock Market Crash, and the stock of existing money inflow begin to decrease. The stock market crash fell for at least six months. The stock market crash fell by at least fifty-five percent. Most of the stock market crash fell by over seventy-percent. The stock market crash down time is inversely proportional to the magnitude of the decline. If the down time is short, the magnitude of the decline is large. If the down time is long, the magnitude of the decline is small. The stock market crash is great harm to investors.

Keywords: M/G/1 queue; Existing Money flow; Stock Crash

1. Introduction

A crash is abbreviation for disaster or the stock market crash. It refers to that because of the influence of random factors, the inherent contradictions of stock market accumulated to a certain extent, then a sudden outbreak of tumbled shares, which will cause abnormal economic phenomenon and bring huge losses to society. Crash is different from the general stock market volatility, but also different from the general stock market risk. In the history, there are some famous crashes happened, such as: the Hong Kong stock market crash in the 1970's, the U.S. stock market crash in 1929 (time: July 1, 1929 - August 1, 1932), Japan's bubble burst in 1989 (time: in December, 1989 - April 2003), Taiwan's crash in the early 1990's, and China's A shares crash in 1993 (time: 1993 February - July 1994) *etc.*

Many researchers study fluctuations of stock price from many aspects. Dr. Guangxi Cao and Minjia Zhang focused on the comparative analysis of extreme values in the Chinese and American stock markets based on the detrended fluctuation analysis (DFA) algorithm using the daily data of Shanghai composite index and Dow Jones Industrial Average. They also studied the difference between the two country's markets especially in 2008. Their studies show that extreme events have nothing to do with the cross-correlation between the Chinese and American stock markets [1].

Researcher Ccile Carpentier and Jean-Marc Suret studied the reflection from the market to the outside environment, including the influence of the external environment for

investors and stock price. Severe external environment will cause losses to a stock listed company, but the influence of institutional investors is weak [2].

Most of scholars study the rule of stock price movement form different aspects [2-11]. Dr. Zhanyou Ma introduced the multiple adaptive vacation policy and the general decrementing service rules based on the classical M/G/1 queuing systems, and obtain the P.G.F. (Probability Generating Function) of stationary queue length by using the embedded Markov chain method and regeneration cycle approach. He put forward the theory about capital stock, laid the foundation of cash flow model [6].

However, their research did not reflect the law of the crash, and cash flow characteristics [1-16]. For example, Dr. Zhanyou Ma [6] just focused on theoretical model research, not the actual application. Dr. Ying Liu developed a Genetic Algorithms to forecast American Shares Price Index. They proposed a prediction method in the future [9].

Data mining is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The goal of the data mining is to extract information from data and transform it into an understandable structure for further use.

Therefore, we propose a method that associate M/G/1 queue system with multiple vacations and server close-down time. Through analyzing large data set in the Shanghai Composite Index from 1990 to now, we want to find out the cause of the crash in the Chinese stock market.

The structure of this paper would appear like this: Section 1 is the introduction of stock research; section 2 presents the model description and embedded Markov chain; Section 3 presents Stochastic Decomposition of Stationary Queue Length and Waiting Time; Section 4 presents the empirical analysis of the money flow in Chinese Stock Market Crash; Section 5 presents the summary and conclusions.

2. Existing Money Flow Model Description

Let $Q_b^{(n)}$ be the number of Commissioned to buy stocks in the system at the completion instant of the nth vacation, then $\{Q_b^{(n)}, n \geq 1\}$ is a Markov chain, and its transition probabilities are given as follows[6]:

$$P_{jk} = \begin{cases} v_{k-j+M}, & k \geq j - M > 0 \\ (1 - H\,((v^*(\lambda)))v_k, & j \leq M, k \neq 1 \\ (1 - H\,((v^*(\lambda)))v_k + H\,((v^*(\lambda)), & j \leq M, k = 1 \\ 0, & j > M, k < j - M \end{cases} \tag{1}$$

where

$$v_j = \int_0^\infty \frac{(\lambda x)^j}{j!} e^{-\lambda x} dV(x), \quad j \geq 0.$$

At the same time, let $\{q_k, k \geq 0\}$ be the steady state distribution of Markov chain $\{Q_b^{(n)}, n \geq 1\}$, i.e.

$$q_k = \lim_{n \to \infty} P(Q_b^{(n)} = k), \quad k \geq 0.$$

$$Q_b(z) = \frac{1}{v^*(\lambda(1-z)) - z^M} \times (v^*(\lambda(1-z))Q_M(z) - ((1 - H(v^*(\lambda)))v^*(\lambda(1-z))$$

$$+ H(v^*(\lambda))z)z^M Q_M(1)). \tag{2}$$

We can get the values of $q_0, q_1, \ldots, q_M$ by using Rouche theorem and Lagrange theorem [12, 13], and they are the queue that is entrusted to buy and sell stocks. Then we can deduce the stochastic decomposition of the stationary performance measures in the steady state system.

3. Existing Money Flow Algorithm with L and W Queue

Theorem 1 If $\rho < 1$ and $\lambda E(V) < M$, the stationary queue length L_v in M/G/1 (GD, MAV) queue can be decomposed into three independent random variables:

$$L_v = L + L_d + L_r ,$$

Where L is the stationary queue length in the classical M/G/1 queue, and the P.G.F. of L can be seen in [14], and the P.G.F. of the additional queue length L_d and L_r are given by [6]

$$L_d(z) = \frac{1 - H(v^*(\lambda))z - \dfrac{1 - H(v^*(\lambda))}{1 - v^*(\lambda)}(v^*(\lambda(1-z)) - v^*(\lambda))}{\left(H(v^*(\lambda)) + \dfrac{1 - H(v^*(\lambda))}{1 - v^*(\lambda)}\lambda E(V) \right)(1 - z)} ,$$

$$L_r(z) = \left(\frac{\beta}{(v^*(\lambda(1-z)) - z^M)(1 - v^*(\lambda(1-z)) + H(v^*(\lambda))(v^*(\lambda(1-z)) - (1 - v^*(\lambda))z - v^*(\lambda)))} \right.$$

$$\left. \times (Q_M(z)(1 - v^*(\lambda(1-z))) - Q_M(1)((1 - v^*(\lambda(1-z)))z^M + H(v^*(\lambda))(z - v^*(\lambda(1-z)))(1 - z^M))) \right) \tag{3}$$

Theorem 2 If $\rho < 1$ and $\lambda E(V) < M$, the stationary waiting time W_v can be decomposed into three independent random variables in M/G/1(GD, MAV) queue,

$$W_v = W + W_d + W_r ,$$

Where W is the stationary waiting time in the classical M/G/1 queue, the LST of W can be seen in [14], and the LST of additional delay W_d and W_r are given by

$$W_d^*(s) = \frac{\lambda - H(v^*(\lambda))(\lambda - s) - \lambda \dfrac{1 - H(v^*(\lambda))}{1 - v^*(\lambda)}(v^*(s) - v^*(\lambda))}{\left(H(v^*(\lambda)) + \dfrac{1 - H(v^*(\lambda))}{1 - v^*(\lambda)}\lambda E(V) \right)s} ,$$

$$W_r^*(s) = \left(\frac{\dfrac{\beta}{\lambda^M v^*(s) - (\lambda - s)^M}}{\lambda(1 - v^*(s)) + H(v^*(\lambda))(\lambda(v^*(s) - v^*(\lambda)) - (1 - v^*(\lambda))(\lambda - s))} \right.$$

$$\left. \times (\lambda^M Q_M(\frac{\lambda - s}{\lambda})(1 - v^*(s)) - Q_M(1)(\lambda(1 - v^*(s))(\lambda - s)^M + H(v^*(\lambda))(\lambda(1 - v^*(s)) - s)(\lambda^M - (\lambda - s)^M))) \right) \tag{4}$$

4. Data Mining and Analysis of Money Flow of Shanghai Composite Index

We analyze the data about Shanghai Composite Index from December 1990 to now. We found four times stock market crash as Table 1. We define two conditions of the stock market crash. Firstly, it dropped by more than fifty percent finally, and continued to decline for more than half a year. Table 1 shows us that Stock Market Crash can drop by seventy percent within half a year. It dropped out seventy-nine percent in maximum. The market began to fall in June 2015, whether it is a new stock market Crash? If it will be a

new stock market crash, it at least will fall to December 2015. It at least will arrive at 2589.

**Table 1. All Previous Stock Market Crash Division of Chinese Stock Market**

Crash umber	The period	Interval values	Fall range
1	1992.5~1992.11	1429.01~386.85	72.9%
2	1993.2~1994.7	1558.95~325.89	79%
3	2001.6~2005.6	2245.43~998.23	55.5%
4	2007.10~2008.10	6124.04~1664.93	72.8%
5	2015.6~?	5178.19~?	>50% ?

We have consulted a large number of references in order to conclude the traditional judgment method of the top of the stock market [1-16]. We found that the most used method is some simple Index, such as Average stock price, Average p / E ratio and Average net rate. The results of the study show that the Shanghai composite index would arrive at the top when p / E ratio reach 60. However, the Shanghai composite index would arrive at the bottom when p / E ratio reach 20. When p / E ratio is over 60, we should sell all the stocks. On the contrary, we should buy stocks when the price is below 20.
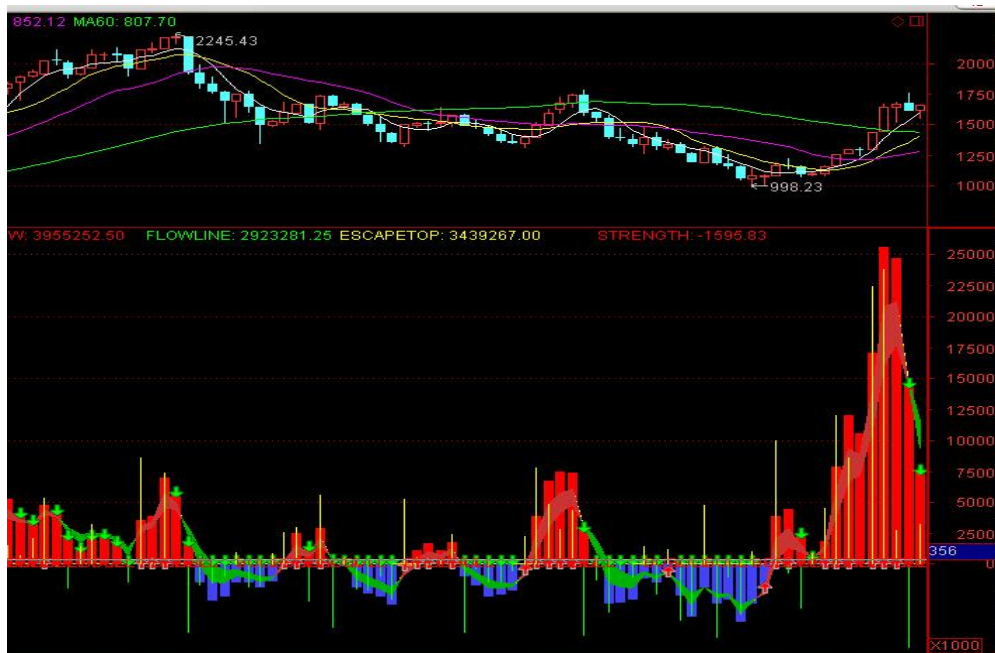

Figure 1. SSE (Shanghai Stock Exchange) Composite Index Monthly Declining Trend Existing Money Flow in Crash Number 1

Figure 1 shows the monthly existing money flow of SSE Composite index between May 1992 and November 1992. In May 1992 the Shanghai composite index arrived at the highest point 1429.01, and this trend continued for a half year. During this period, the monthly existing money flow has continued to flow out. The market was no any resistance, and dropped by seventy-two point nine percent finally. It has consecutive overcast K line. Green block in Figure 1 shows us to be in bear market. The red arrow in June 1992 indicates that the market would be beginning to decline as Red post shorten, and the bull market is end. Stock existing money flow algorithm clearly tells us the trend, and investors can make investment decisions based on our research results.
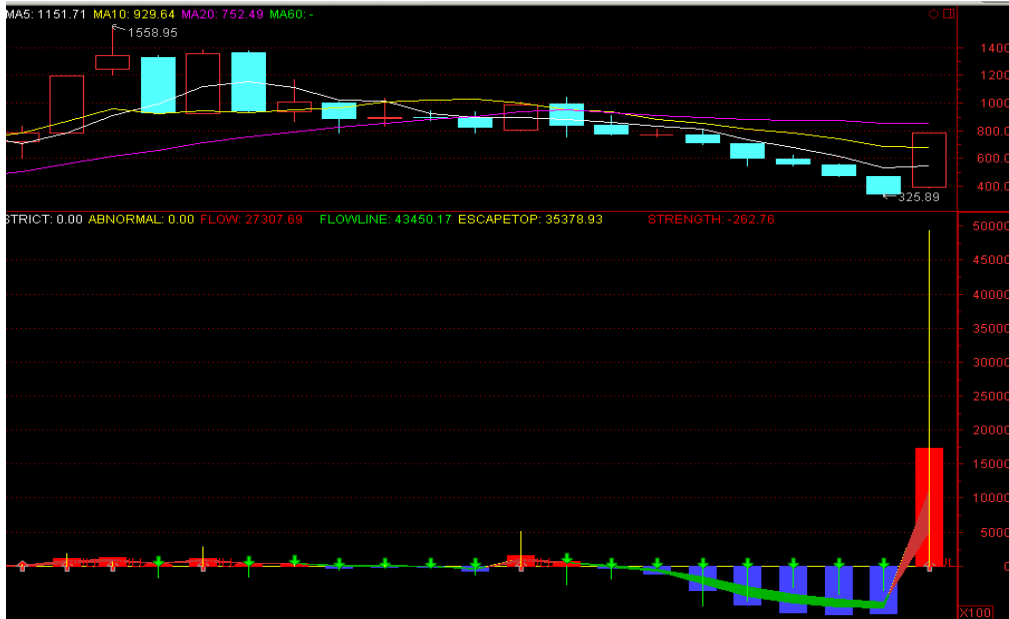
Figure 2. SSE (Shanghai Stock Exchange) Composite Index Monthly Declining Trend Existing Money Flow in Crash Number 2

Figure 2 shows the monthly existing money flow of SSE Composite index between February 1993 and July 1994. In May 1993 the Shanghai composite index arrived at the highest point 1558.95, and this trend continued for 17 months. During this period, the monthly existing money flow has encountered strong resistance for one year. This year has been sideways shocks, and the monthly existing money inflow and outflow is very few. At last markets dropped by seventy-nine point. Especially the last time accelerates the bottom and it has five consecutive overcast K line. In last five months the monthly existing money outflow is also increasing. It arrives at 325.89 in July 1994. It fell more than the first crash.
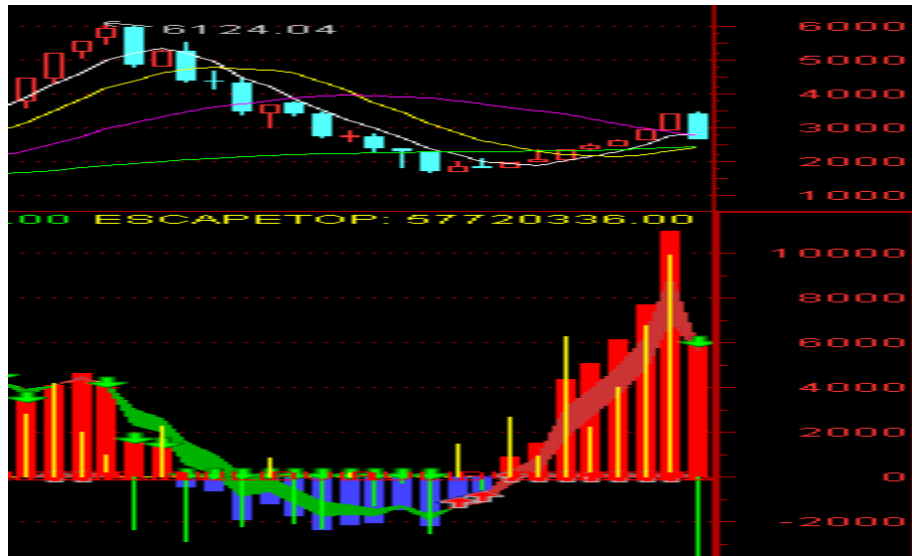

Figure 3. SSE (Shanghai Stock Exchange) Composite Index Monthly Declining Trend Existing Money Flow in Crash Number 3

Figure 3 shows the monthly existing money flow of SSE Composite index between June 2001 and June 2005. In June 2001 the Shanghai composite index arrived at the highest point 2245.43, and this trend continued for four years. During this period, the monthly existing money flow has encountered strong resistance for three years. This period has been sideways shocks, and the monthly existing money inflow and outflow is perfect. During this period we can find many invest opportunity. It arrives at 998.23 in June 2005. It is a bear market with four years; however it just drops by fifty-five point. It is very small bear market.

Figure 4. SSE (Shanghai Stock Exchange) Composite Index Monthly Rising Trend Existing Money Flow with Bull Market in 2007

Figure 4 show the monthly existing money flow of SSE Composite index between June 2005 and October 2007. In June 2005 the Shanghai composite index arrived at the lowest point 998.23, and then continued to rise until October 2007. The highest point is 6124.04. The graph is based on our proposed stock money flow algorithm for data mining, and we get the money flow graphics. All of the data come from the exchange of real data. The computer can draw out money flow graph according to the transaction data in time. The bottom half of the graph is our stock money flow chart. Red columns show money inflows, and Blue columns show money outflows. Red block indicates that the market is in short trend, and Green block indicates that the market is in a bull trend. From Figure 4 we can see that the stock will rise with money inflows, the stock will fall with money outflows. Figure 4 can tell us the round of bull market is the money to push up the rise. The end of the bull market is significantly reduced. Especially Stock price increases, but capital inflows are reduced in October 2007. Stock prices hit a new high point. This is a precursor to the top, and Crash is imminent. If there is the case, the stock market crash is coming. You'd better sell all the stocks to keep your income and principal.

Figure 5. SSE (Shanghai Stock Exchange) Composite Index Monthly Declining Trend Existing Money Flow in Crash Number 4

Figure 5 shows the monthly existing money flow of SSE Composite index between October 2007 and October 2008. In October 2008 the Shanghai composite index arrived at the lowest point 1664.93, and this trend continued for a whole year. During this period, the monthly existing money flow has continued to flow out. The market was almost no any resistance, and dropped by seventy-two point eight percent finally. Green block in Figure 5 shows us to be in bear market. The red arrow in November 2008 indicates that the market would be beginning to rebound, and the bear market is end.

Figure 6. SSE (Shanghai Stock Exchange) Composite Index Monthly Declining Trend Existing Money Flow in Crash Number 5

Figure 6 shows that the monthly money flow of SSE Composite index began from June 2015. The Shanghai composite index arrived at the highest point 5178.19, and ended the bull market for 15 months. The world thinks that Chinese encountered Stock Market Crash. If it is one Stock Market Crash, it will last 6 months, and drop by fifty percent finally. At present our algorithm tells us that it is in bear market. Investors can only wait patiently.

5. Conclusion

From the proofs and the results of the above new money flow model, we do some research in Chinese security market to study Chinese Stock Market Crash. With the proposed method and data mining from the five crashes in Chinese stock market, we find the law of the top of Chinese stock market. At the top of the stock price, money flows begin to decrease, which is the best time for the smart investors to escape. But if you began to escape when money flow outflow coming, you would be too late for sell. Because you may have been locked on the top, then you can only face a loss. In the future, we will study American stock market, European stock market and Japanese stock market to enhance the accuracy of predicting the magnitude and duration of crash down, and the research results show that it will be worth investing.

Acknowledgements

References

[1] G. X. Cao and M. J. Zhang, "Extreme values in the Chinese and American stock markets based on detrended fluctuation analysis", Physical A: Statistical Mechanics and its Applications, vol. 436, no. 1, **(2015)**, pp. 25-35.

[2] C. Carpentier and J. M. Suret, "Stock market and deterrence effect: A mid-run analysis of major environmental and non-environmental accidents", Journal of Environmental Economics and Management, vol. 71, no. 1, **(2015)**, pp. 1-18.

[3] Q. Z. Xu, "A Novel Machine Learning Strategy Based on Two-Dimensional Numerical Models in Financial Engineering", Mathematical Problems in Engineering, vol. 2013, no. 1, **(2013)**, pp. 1-6.

[4] Q. Z. Xu, J. Y. Wu and Q. Chen, "A Novel Mobile Personalized Recommended Method Based on Money Flow Model for Stock Exchange", Mathematical Problems in Engineering, vol. 2014, no. 1, **(2014)**, pp. 1-9.

[5] J. Zhu, Y. Yang, Q. Xie, L.W. Wang and S. Hassan, "Robust Hybrid Name Disambiguation Framework for Large Databases", Scientometrics, vol. 98, no. 3, **(2013)**, pp. 2255-2274.

[6] Z. Y. Ma and Q. Z. Xu, "General decrementing service M/G/1 queue with multiple adaptive vacations", Applied Mathematics and Computation, vol. 204, no. 1, **(2008)**, pp. 478-484.

[7] Q. Z. Xu, "Continuous time M/G/1 queue with multiple vacations and server close-down time", Journal of Computational Information Systems, vol. 3, no. 2, **(2007)**, pp. 753-757.

[8] Q. Z. Xu and Z. Y. Ma, "Discrete time Geo/G/1 queue with Bernoulli gated service simulation system", Applied Mathematics and Computation, vol. 204, no. 1, **(2008)**, pp. 37-44.

[9] Q. Z. Xu and Y. Liu, "A Genetic Algorithms to forecast American Shares Price Index", Journal of Information and Computational Science, vol. 8, no. 8, **(2011)**, pp. 1245-1250.

[10] Z. Y. Ma, "The Geom/G/1 queue with multiple vacation and server set-up/close times", Operation Research and Management Science, vol. 13, no. 5, **(2004)**, pp. 21-25.

[11] J. Huij and J. Derwall, "Global equity fund performance, portfolio concentration, and the fundamental law of active management", Journal of Banking & Finance, vol. 35, no. 1, **(2011)**, pp. 155-165.

[12] E. C. Tichmarsh, "Theory of Functions", Oxford University Press, London, **(1952)**.

[13] E. T. Whittaker and G. N. Watson, "A Course of Modern Analysis", Cambridge University Press, London, **(1946)**.

[14] H. Takagi, "Queuing Analysis: Vacation and Priority Systems", Elsevier Science Publishers, **(1991)**, vol. 1, pp. 202-227.

[15] P. Kanniappan, "Necessary Conditions for Optimality of No Differentiable Convex Multi-objective Programming", Journal of Optimization Theory and Applications, vol. 40, no. 2, **(1983)**, pp. 167-174.

[16] T. Amahroq and N. Gadhi, "Second order optimality conditions for the extremely problem under inclusion constraints", Journal of Mathematical Analysis and Applications, vol. 285, no. 1, **(2003)**, pp. 74-85.