Application of XCSR Model for Dynamic Portfolio Selection

Mei-Chih Chen
Minghsin University of Science and Technology
National Chiao Tung University
E-Mail: mindy.must@gmail.com

Chang-Li Lin
National Chiao Tung University
E-Mail: changli.lin@gmail.com

An-Pin Chen
National Chiao Tung University
E-Mail: apc@iim.nctu.edu.tw

ABSTRACT

Security selection is the most time-consuming problem in investment process. Today, investing environment is more complex than before and investors can't see through the information frame, the reality behind which they are unable to scrutinize. Static model cannot provide appropriate solutions when current phenomena are completely different from that of the training period. Learning classifier system (LCS) is possessing with dynamically learning mechanism to evolve internal rules in response to changes in the external environment. This study employs real number version classifier system (XCSR) to investigate constructing dynamic stock portfolio. We examine the constituents of the TSEC Taiwan 50 Index and take those stocks daily transaction data to generate a number of technical indicators as input factors for XCSR model. The empirical results revealed that this study's XCSR model yields return on investment is significantly better than the Buy and Hold model. This research also indicates that classifier system can effectively monitor market fluctuations and the proposed model can help investors obtain relatively optimal returns.

Keywords: Extended Classifier System, Dynamic Portfolio, Framing Effect

INTRODUCTION

Behavioral finance researchers found that the securities individual investors buy subsequently underperform those they sold (Odean,1999) and also discovered that the more people traded, the worse they did (Barber and Odean,2001). Furthermore, the framing effect of prospect theory (Kahneman and Tversky, 1979) indicated that the tendency of decision making based on the context in which a choice is presented. Under circumstances of limited knowledge, investors are invariably reactions directly to the received information, the reality behind which they are unable to scrutinize (Kuo and Chiang, 2004).

Stock market investment requires making decisions under uncertain situation in investment period which consists of strategic asset allocation, tatical asset allocation, and stock picking (Amenc and Le Sourd, 2003). Strategic asset allocation is subject to long-term assembling an asset level allocation. Tatical asset allocation is short-term allocation in order to take advantage of market opportunities. Selection stocks is the most time-consuming stage (Amenc and Le Sourd, 2003) and a difficult task, but for which people are most overconfident (Barber and Odean, 2001). The performance of individual security has a greater impact on the return of portfolio (Hensel, Ezra and Ilkiw, 1991). Financial market is a complex and continuously changing environment. Individual investors and sometimes even professional fund managers allow their emotions to get in the way of rational investment decision-making (Beach and Rose, 2005). Therefore, our study focused on tatical allocation and stock selection of portfolio to prevent investors from framing effect, and reduce common investment mistakes.

Many types of security selection methods using artificial intelligence have been proposed (Chan et al, 2002; Venugopal, Subramanian and Rao, 2004) in recent years, and the use of artificial intelligence for financial market trend analysis and forecasting has been increasing as artificial intelligence gradually comes into favor again. Nevertheless, such neural networks, decision trees, use historical data for learning and training to produce a fixed prediction model. But Historical data don't constitute a good representation of the forthcoming period (Amenc and Le Sourd, 2003). Therefore, this type of forecasting model cannot perform well when the real environment is completely different from the past. For their part, classifier systems absorbed evolutionary computing and reinforcement learning mechanisms based on their dynamic environmental learning concept. These features enhance system accuracy and performance. Thus this study regards the financial market as a continuously changing environment, and consequently employs an XCS in

conjunction with a real number encoding scheme to construct a dynamic stock portfolio selection model (XCSR model). We also perform empirical testing to verify the performance of the model for weekly and biweekly adjustment cycle, using the Sharpe ratio and return on investment as performance assessment criteria. Eventually, the proposed model performance is compared with that of Buy and Hold model.

INVESTMENT PORTFOLIO AND LEARNING CLASSIFIER SYSTEM

The concept behind investment portfolios is to combine several different investment targets to avoid concentrating too much risk on any one target with the aim of dispersing overall investment risk. Over a half century, the Markowitz mean-variance model has become a universally understood technique within the investment field. However, this model is limited by the uncertainty of the inputs such as expected returns, standard deviations, and correlation matrix. Which assumptions are not consistent with the real environment (Pawley, 2005) and consequently portfolios created using this method rarely yield significantly positive performance. Jahnke (1997) pointed asset allocation should be viewed as a dynamic process. It should take into consideration both pension obligation and capital market opportunities, including risk, otherwise makes no economic sense. The portfolio optimization and selection is a complex task (Venugopal, Subramanian and Rao, 2004) because there are a wide range and variety to choose from.

The use of information technology for investment portfolio has generally focused on the two aspects of investment target selection and optimal asset proportion management. For instance, Chan et al. (2002) proposed a fuzzy rule-base stock selection model with rate of return, current ratio, and yield rate as input factors. Their model uses Genetic algorithm to find each company's appraisal grade and employs a multi-period random capital allocation model; empirical results indicate that investment portfolios constructed using this method perform well in terms of predicted rate of return, variance, and utility value. Venugopal, Subramanian and Rao (2004) proposed a Genetic Algorithm Model for portfolio selection. The model considers both equity and debt securities and vice versa. The computerized dynamic portfolio has outperformed the Sensex (Bombay Stock Exchange Sensitive Index) throughout the testing period.

In recent years, classifier system techniques have been used in many different fields, including data mining (Barry, Holmes and Llora, 2004), electrical machinery control systems (Carse , Fogarty and Munro, 1996), and financial market analysis, and have demonstrated excellent performance in all of these areas. Classifier systems were

introduced by Holland in the form of cognitive systems. The preliminary learning classifier system (LCS) framework was introduced in 1977 (Holland and Reitman, 1977). Wilson (1995) proposed an extended classifier system (XCS) following continuous improvement by many researchers. Beltrametti et al. (1997) used an LCS model to study the foreign exchange market, the empirical results of this research showed that classifier systems can classify external information and generate suitable predictions for changing environment. In addition, other scholars have used classifier systems to analyze the trading of individual stocks. For instance, Liao and Chen (2001) and Schulenburg and Ross (2002); both obtained experimental results significantly better than Buy and Hold and random trading strategies respectively.

THE PROPOSED MODEL

This study's research model is shown in Figure. 1. Data on the constituent stocks of the TSEC Taiwan 50 Index were taken from a historical stock trading database; all constituent stock was subjected to pre-processing and then submitted to an XCSR model. Each XCSR model forecasts the future return of one stock in accordance with the adjustment cycle. All the investment targets forecast to yield good returns were then assembled as a portfolio which will be invested in the succeeding cycle. For example, the adjustment cycle is n days, then all capital must be divided into n equal portions at each cycle; one portion of capital is invested on each day, and capital are averagely invested in the recommended investment targets generated by XCSR model. On the second cycle, the n+1th day, all positions established on the first day of previous cycle are sold at the opening price, and the capital obtained in this way is evenly spread across the recommended investment targets and are purchased at the opening prices. This method is continued until the end of the investment period. Any cash remaining is used to purchase odd lots and after trading is used during the next trading. While the Buy and Hold model's adjustment cycle and capital allocation were the same as in the XCSR selection model, the portfolio was assembled by all the constituent stocks of the TSEC Taiwan 50 Index without any selection processing.

Research Target and Data Pre-processing

The research target of this study consisted of the constituent stocks of the Taiwan 50 Index as of April 7, 2006. These stocks comprised roughly 70% of the aggregate market value of the Taiwan 50 Index at that time, and comprised an even higher 0.989 of the TSEC weighted stock index linked correlation index http://www.tw50etf.com/tw50etf/tw50/Introduction/). The stocks chosen as the target of this study are

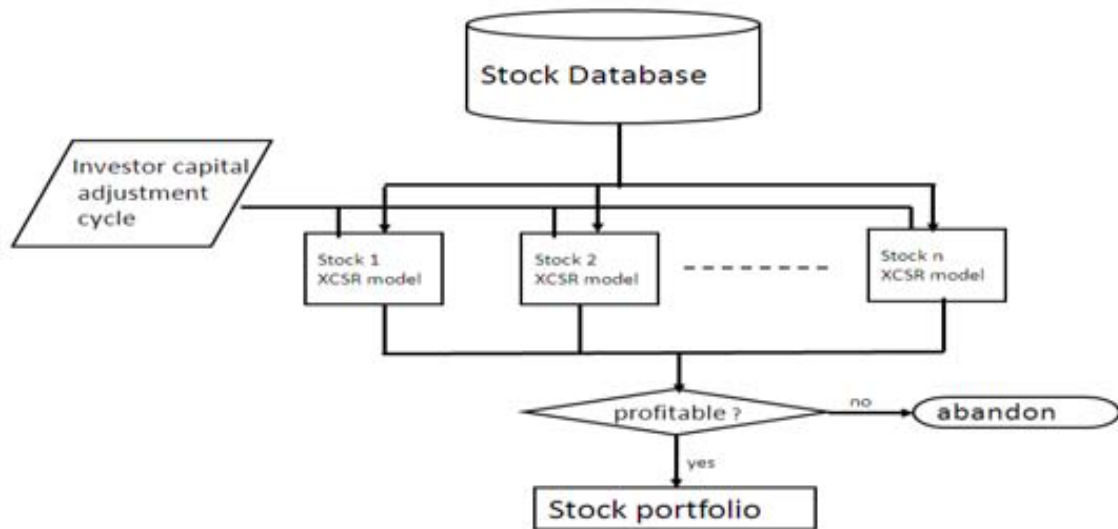consequently highly representative of the market.

Figure 1 XCSR stock selection model

There are five technical indicators used in this study. We converted daily raw opening price, closing price, maximum price, minimum price, and trading volume into five technical indicators moving average (MA), stochastic indicators (KD), moving average convergence divergence (MACD), relative strength index (RSI) and Williams %R (WMS %R) which are usually used by investors. Because the indicators have different numerical ranges, we normalized the five indicators in the range of (0,1) using min-max normalization .

We also calculated the proportional increase or decrease percentage in each the five indicators between two trading days as a strength correlation input factor expressing the strength of correlation between market changes on different days.

The proportional increase or decrease is calculated using formula (1):

$$r_t = \frac{x_t - x_{t-1}}{x_{t-1}} \cdot (1)$$

Where

xt is the technical indicator value on day t

xt-1 is the technical indicator value on day t-1.

XCSR Selection Model

This study sets parameters using the optimal values recommended by Wilson (1995). The genetic algorithm acts on the match set of classifier system also adopts the recommend parameters by Wilson (1995). Additionally, this study employs Crossover algorithm and Mutate algorithm proposed by Loiacono's (2004).

The structure of a classifier can be divided into condition and action parts. The condition part includes the five technical indicators MA, KD, MACD, RSI, and WMS %R, and the proportional increase or decrease percentage in the five indicators relative to the previous day. The action part contains the forecast {rise: 1, fall: 0} reflecting price fluctuations n days (the adjustment cycle) in the future.

A classifier system with binary encoding uses a string composed of {0,1,#} to express the state of the real-world environment. Sometimes a binary variable is taken to represent a pre-threshold continuous one, but then the threshold has not been done adequately (Wilson, 2000). Thus this study uses a real number encoding method in order to accurately describe the environment states. Center-spread representation (Wilson, 2000) is employed to encode input real numbers.

Classifier systems are rule-based systems. Each classifier has its fitness strength that shows its usefulness in the current system. After a classifier has been chosen, it receives a reward in case of successful prediction; otherwise, it pays compensation for incorrect prediction. In this research, the reward allocation based on prediction accuracy (Wilson, 1995) is shown in Table 1.

Table 1 Reward allocation

Classifier forecasting	Market condition	Reward
rise	rise and > transaction cost	+ reward
rise	rise but <= transaction cost	no reward
rise	fall	- reward
fall	fall	+ reward
fall	rise	- reward

EMPIRICAL RESULT

The empirical part of this study used TSEC data. The trading date, opening price, closing price, maximum price, minimum price, trading volume, and trading value of each of the constituent stocks were extracted from daily trading data.

The testing process used June 30, 2003 as a dividing date, and performed model training using all daily data from the date each constituent stock was listed to June 30, 2003. The rule sets established during the training period are used during the testing period as initial rules sets. Testing was performed using daily data from July 1, 2003 to February 27, 2007. A total of 910 data sets were used. The initial investment was NT$10 million. There wasn't short selling of the investment targets, and subscription, redemption, conversion, or dividend activity during the investment period. Portfolio total values are calculated to the day's closing prices. The cost of each trade was taken into consideration. Service fees were 1.425 thousandths and securities trading tax was 3 thousandths.

Since the cash obtained from each trading session were used for continued trading, the return on investment was expressed as the annual compound interest rate r, which was calculated as shown in formula (2):

$$E = B \times (1 + r)^n \quad \Rightarrow \quad r = \sqrt[n]{\frac{E}{B}} - 1 \quad (2)$$

E is the final total value
B is the initial amount invested
r is the interest rate obtained
n is the number of periods, expressed in years

The Sharpe ratio was used to calculate profitability per unit risk; the risk-free interest rate was set as the post office two-year CDs (certificate of deposit) annual rate of 2.185%. This empirical used both explore and exploit models for XCSR rule selection mechanisms. Explore and exploit are used alternately during the training period to prevent the fitness of a specific classifier from being excessively high and preventing other classifiers from being selected and implemented, which would affect overall system accuracy (Butz and Wilson, 2002). Exploit was used as the selection mechanism during the testing period.

The testing results (Table 2) indicate that the proposed XCSR model yields much better performance than the Buy and Hold model. The prediction accuracy of XCSR selection model is shown in Table 3. The worst and best prediction probability of XCSR is only a slightly different. Obviously, the proposed XCSR model is really functional and ability of prediction is also stable.

Table 2 Comparison of Testing Results

Adjustment cycle	ROI (year)		Sharpe ratio (year)	
	XCSR	Buy and Hold	XCSR	Buy and Hold
Weekly	44.69%	-8.32%	215.45%	-79.83%
Biweekly	28.30%	-2.05%	141.61%	-20.47%

Table 3 Prediction Accuracy of XCSR Model

XCSR prediction accuracy probability			
Adjustment cycle	Average (%)	Worst(%)	Best (%)
weekly	62.37	61.93	62.68
biweekly	65.11	64.81	65.62

CONCLUSION

Classifier system is an on-line learning system and reinforcement from environment based on an evolving set of classifiers (Wilson, 2000). Useful classifiers gain strong fitness are selected and propagated over others less useful, thus the system performance increase gradually. Therefore an XCSR model which is an XCS model together with real number encoding was employed in this study. The input factors consisted of MA, KD, MACD, RSI, and WMS %R and an empirical testing was performed by assuming weekly and biweekly adjustment cycles and comparing the results of the model with Buy and Hold model. Testing results showed that the proposed model successfully to keep track of market trends. Regardless of whether the user selects a weekly or biweekly adjustment cycle, the XCSR portfolio selection model yields return on investment and Sharpe ratio better than those of the Buy and Hold model.

Risk management is another important issue. Future research should consider incorporating estimated risk values and assess possible portfolio risk. Researchers may further investigate the input factors and select appropriate indicators for different investment periods and different asset so as to increase profitability. As for trading strategies, stop-loss and stop-profit mechanisms can be used to avoid unnecessary trading costs. Finally, portfolio insurance policy could be taken into consideration to get free from system risk (Chen, Huang and Chen, 2006).

REFERENCES

Amenc, N. and Le Sourd, V. (2003). *Portfolio theory and performance analysis.* Chichester, England: John Wiley and Sons.

Barber, B.M. and Odean, T. (2001). Boys will be boys: gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, 116, 261-292.

Barry, A., Holmes, J. and Llora, X. (2004). Data mining using learning classifier systems. Applications of Learning Classifier Systems. *Springer-Verlag*, 150, 15-60.

Beach, S.L. and Rose, C.C. (2005). Does Portfolio Rebalancing Help Investors Avoid Common Mistakes?. *Journal of Financial Planning*, 18(5), 56-61.

Beltrametti, L., Fiorentini, R., Marengo, L. and Ramborini. (1997). A learning-to-forecast experiment on the foreign exchange market with a classifier system. *Journal of Economic Dynamics and Control*, 21(8), 1543-1575.

Butz, M.V. and Wilson, S.W. (2002). An algorithmic description of XCS. *Soft Computing-A Fusion of Foundation, Methodologies and Applications*, 6(3-4), 143-283.

Carse, B., Fogarty, T.C. and Munro, A. (1996). Evolving fuzzy rule based controllers using genetic algorithms. *Fuzzy Sets and Systems*, 80, 273-294.

Chan, M.C., Wong, C.C., Tse, W.F., Cheung, Bernard K.-S. and Tang, Gordon Y.-N. (2002). Artificial intelligence in portfolio management. *Lecture Notes in Computer Science*, 2412, 159-166.

Chen, M.C., Huang M.C. and Chen, A.P. (2006 , July 12-14). *A XCS-Based approach for portfolio insurance policy*. The International Conference on Business and Information, Singapore.

Hensel, C.R., Ezra, D.D. and Ilkiw J.H. (1991). The Importance of the asset allocation decision. *Financial Analysts Journal*, 47(4), 65-72.

Holland, J. H. and Reitman, J. S. (1977). Cognitive systems based on adaptive algorithms. *ACM SIGART Bulletin*, (63), 49-49.

Jahnke, W. W. (1997). The Asset Allocation Hoax. *Journal of Financial Planning*, 10(1), 109-113.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometric*, 47(2), 263-291.

Kuo, M.H. and Chiang, N.J. (2004). The Framing Effect : Does the Market see through the Announcement of Insider Shoring?. *Web Journal of Chinese Management Review*, 7(2), 30-45.

Liao, P.Y. and Chen, J.S. (2001). Dynamic trading strategy learning model using

learning classifier systems. *Proceedings of the 2001 Congress on Evolutionary Computation* ,2 , 783-789.

Loiacono, D. (2004). *Evolving rules with XCSF: Analysis of generalization and performance*. University of Politecnico di Milano, Italy: Unpublished master's thesis.

Odean, T. (1999). Do Investors Trade Too Much?. *The American Economic Review*, 89(5), 1279-1298

Pawley, M.G. (2005, September). *Resampled Mean-Variance optimization and the dynamic nature of markets*. The biennial conference of the Economic Society of South Africa, Durban.

Schulenburg, S. and Ross P. (2002). Explorations in LCS models of stock trading. *Lecture Notes in Artificial Intelligence*, 2321, 150-179.

Venugopal, M.S., Subramanian, S. and Rao, U.S. (2004). Usefulness of genetic algorithm model for dynamic portfolio selection. *Journal of Financial Management & Analysis*, 17(1), 45-53.

Wilson, S.W. (1995). Classifier fitness based on accuracy. *Evolutionary Computation*, 3(2), 149-175.

Wilson, S.W. (2000). Get real! XCS with continuous-valued inputs. *Lecture Notes in Computer Science*, 1813, 209-219.