Expert Systems with Applications 38 (2011) 151–154

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

# Using the XCS classifier system for portfolio allocation of MSCI index component stocks

Wen-Chih Tsai *, An-Pin Chen

*Institute of Information Management, National Chiao Tung University, Taiwan*

ARTICLE INFO

*Keywords:*
Financial forecasting
XCS
MSCI Taiwan index component stock
Reinforcement learning

ABSTRACT

In a recent study, Schulenburg and Ross (2001) proposed the LCS for short-term stock forecast. Studley and Bull (2007) proposed the extended classifier system (XCS) agent to model different traders by supplying different input information. Announcement made by Morgan Stanley Capital Investment (MSCI) regarding the additions, removals, and even the weights of the component stocks in its country indices every quarter generally would cause changes to the prices and/or trade volumes of the associated component stocks. This paper takes an XCS in artificial intelligence to dynamically learn and adapt to the changes to the component stocks in order to optimize portfolio allocation of the component stocks. Since these price trends of MSCI component stocks are influenced by unknown and unpredictable surroundings, using XCS to model the fluctuations on financial market allows for the capability to discover the patterns of future trends. This simulation works on the basis of the changes to 121 component stocks in the MSCI Taiwan index between 1998 and 2009 suggests the XCS can produce the great profit and optimize portfolio allocation.


## 1. Introduction

Stocks that are added to major stock indices usually see rises in their trading volumes and prices after the announcements. On the other hand, investors usually sell the stocks after their removals from MSCI indices. This phenomenon has been widely studied in finance Kovalerchuk and Vityaev (2000), Liao and Chen (2001) and Chen and Chen (2006). Most studies to date have focused on US stocks. This paper turns attention to one of the emerging markets. This paper documents and analyzes changes to the returns and the trade volumes in response to additions and removals of the component stocks in MSCI country indices.

Previous works approached the problem in a variety of different ways. They were usually about the institutional investment by means of the traditional statistical analysis and financial engineering. And (McIvor, McCloskey, Humphreys, & Maguire, 2004) they all shared some common characteristics such as technical indicators, quantizing financial series, static prediction rule, and buy or sell prediction signal.

Chen and Chen (2006), Tsai and Chen (2008), Schulenburg and Ross (2001), etc. have applied learning classifier system (LCS) to different financial systems such as the future market, the foreign exchanges market, the derivatives market and the equity market. While their results were promising, it is widely accepted that better results can be achieved by improving the processes surrounding the main XCS learning components. The following are reasons to use XCS on dynamic and noisy environments:

- XCS is able to evaluate rules that are ideal for modeling problems without retraining all data.
- XCS has been shown to properly learn from noisy, complex, and non-linear environments when the outside information continuously changes.
- XCS is capable of making real-time and accurate learning and responses.
- XCS can discover generally accurate rules to perform on a variety of problem domains (Wilson, 1996).
- XCS can adjust itself to strengthen its inward knowledge step by step.

The stock market traders (Hashemi, Blanc, Rucks, & Rajaratnam, 1998) always look for trading strategies to optimize portfolio allocation and to obtain high returns. This paper presents empirical results for the XCS agents trading system providing trading strategies for the Taiwan MSCI component stocks.

## 2. Literatures review

### 2.1. LCS

The classifier systems have been gaining more and more popularity in the artificial intelligence (AI) domain. Holland (1975) and

* Corresponding author. Tel.: +886 911296906.
*E-mail addresses:* miktsai@gmail.com, ibm@iim.nctu.edu.tw, wctsaie@tsmc.com (W.-C. Tsai), apc@iim.nctu.edu.tw (A.-P. Chen).

0957-4174/$ - see front matter 
doi:10.1016/j.eswa.2010.06.031

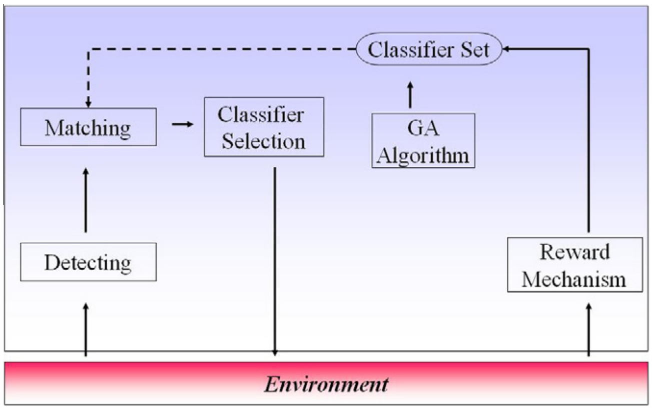


**Fig. 1.** System architecture of LCS.

Holland and Reitman (1978) introduced the concept the mechanism of learning classifier systems (LCS) in the 1980s, and a series of researches has focused on the problem derived from LCS, such as generalization, the classifier syntax, the credit allocation procedure, the discovery component, and the internal messages list in Holmes (1996). The schematic overview of LCS in Fig. 1 shows three major functional components of LCS:

- The detecting function which allow the machine to interact with its environment.
- Reward mechanism allows separating successful rules from unsuccessful or meaningless rules.
- Genetic algorithm (Glodberg, 1989; Yuan & Shaw, 1995; Yuan and Zhung, 1996) used to generate more optimal rules for the rule set.

### 2.2. XCS

Recently, the extended classifier systems (XCS) have been becoming the primary research topic in AI, especially in the financial domain (Chen & Chen, 2006; Studley & Bull, 2007; Tsai & Chen, 2008). The schematic overview of XCS in Fig. 2 provides three major functional phases of XCS:

- Transaction data encoding phase decodes and normalizes the different stocks.
- Knowledge extraction phase allows separating successful rules from unsuccessful or meaningless rules.
- Knowledge integration phase generates more optimal rules including the single stock perdition rule and portfolio allocation rule.

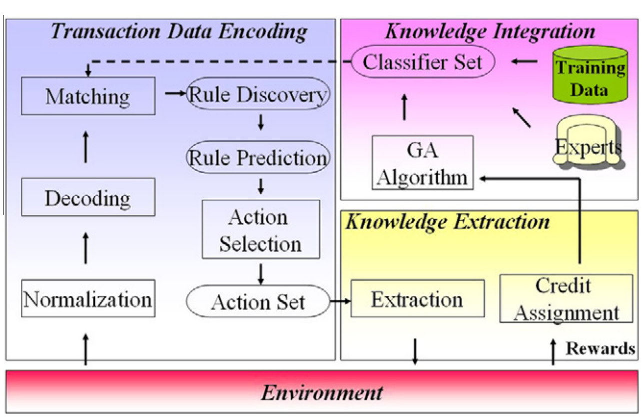


**Fig. 2.** System architecture of XCS.

## 3. System architecture

The key quantity the trading system predicts is the percentage of current total investment in MSCI component stocks. This is closely related to the expected changes to the prices in the subsequent trading days Wang (2003). Three major phases in Fig. 2 allow XCS agents to predict future price movements based on different sets of historical financial data.

### 3.1. Algorithm

The XCS algorithm is shown in Table 1. The XCS algorithm describes how to forecast from the input to result.

### 3.2. Transaction and encoding phase

This phase generally normalizes the MSCI component stock prices and trading volumes to the status such as up, down or ignored. The decoding sample is given in Table 2.

### 3.3. Knowledge extraction phase

During this phase, the XCS trading system follows the action set to adjust the portfolio such as increasing the percentages of 2303.tw and 2408.tw while reducing the percentages of 2330.tw and 2345.tw. The action set table is given in Table 3.

**Table 1**
Algorithm of XCS.

```
XCS Algorithm
Input RSq<- q rule sets(RS) from different knowledge sources
Output one integrated rule set
procedure XCS
  Initialize classifier set
  While (termination condition of XCS is false)
    Get environment state
    Normalized the state
    Decode the state
    Generate match set
    Generate prediction rule
    Selection action
    Generate action set
    do winner action
      Get rewards
      Update attribute-values of relevant classifiers
      trigger Genetic Algorithm
        Selection
        Crossover
        Mutation
      end trigger
    end do
  end while
  Report the execution and learning performances
  Store the learned classifier set
end procedure
```

**Table 2**
Decoding.

| Date | Symbol | High | Low | Price | Volume | Change | Decode |
|---|---|---|---|---|---|---|---|
| 31-March-09 | 2330.tw | 0 | 0 | 0 | 0 | 0 | 00000 |
| 30-March-09 | 2330.tw | 1 | 1 | # | 1 | # | 11#1# |
| 27-March-09 | 2330.tw | 1 | 1 | 1 | 1 | 1 | 11111 |
| 26-March-09 | 2330.tw | 0 | 1 | 1 | 0 | 1 | 01101 |
| 25-March-09 | 2330.tw | 1 | # | # | 1 | 1 | 1##11 |

**Table 3**
Action set.

| Company ID | Name | Action | As-Is (%) | To-Be (%) |
|---|---|---|---|---|
| 2330.tw | TSMC | Sell | 23 | 20 |
| 2303.tw | UMC | Buy | 9 | 10 |
| 2408.tw | ⋮ | Buy | 3 | 5 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 2345.tw | Accton | Sell | 3 | 0 |

**Table 4**
Action set.

| Factor code | Action | Currency (%) |
|---|---|---|
| 11#110001#10 | Sell | 82 |
| 01111#0#0110 | Buy | 81 |
| 01#101#10001 | Sell | 79 |
| 01011#0#1101 | Buy | 79 |
| 01#111#10000 | Sell | 73 |
| ⋮ | ⋮ | ⋮ |
| 101##0011101 | Buy | 53 |

*3.4. Knowledge integration phase*

In this phase, the XCS trading generates more optimal rules including the single stock perdition rule and portfolio allocation rule. The rule set is shown in Table 4.

## 4. Simulation

*4.1. Simulation*

Experiment is performed on the MSCI Taiwan index component stock data, which is derived from Securities & Futures Institute (SFI). The trading period for this experiment is from January 1998 to March 2009. Since the simulation corresponds to the foreign capital, the stocks which below to the MSCI Taiwan index component. The condition part of the classifiers is within 5 days; while the action part is the trend of the next 3 days. There are some preprocess to be completed before having the data placed into XCS.

- Missing values will be deleted due to the concept of data cleaning.
- The five attributes, which are buy/sell of QFII, buy/sell of securities investment trust, balance of margin purchasing, balance of short selling, and volume, are the most important factors when it comes to the institutional analysis. Thus, they are taken into account in the simulation. Each condition will be considered into four parts on the basis of discretion technique.
- The action part will also be described into four regions corresponding to levels of the upward or downward trends for the next trading day. The format of classifiers is shown in Table 4, and the simulation procedures are shown in Fig. 3.

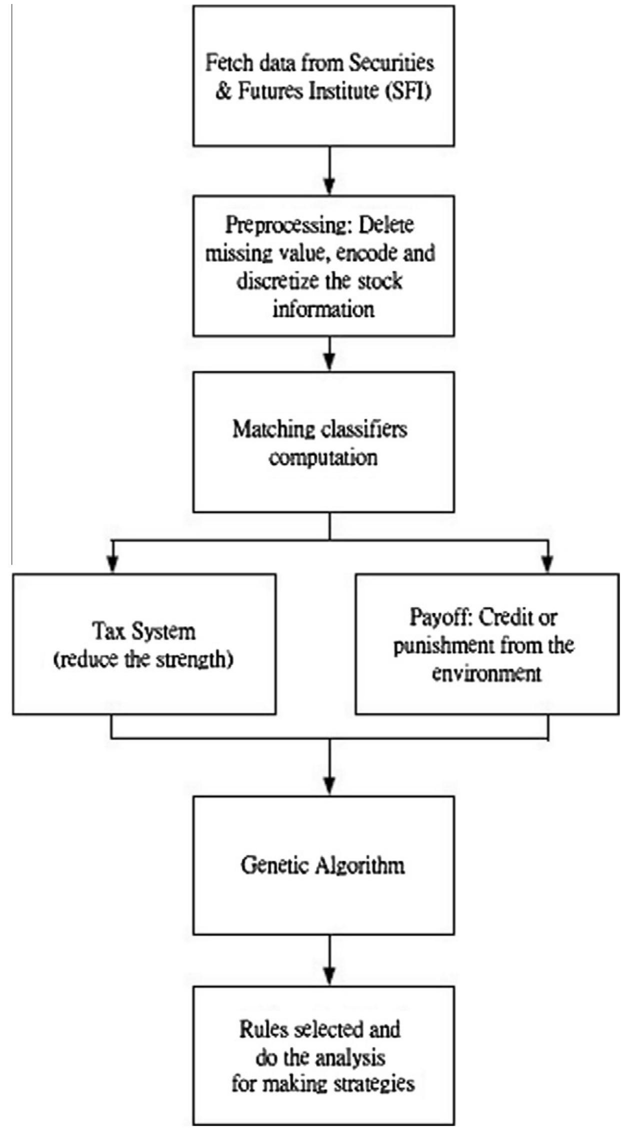


**Fig. 3.** Simulation procedure.

*4.2. Implementation*

Due to the historical data of MSCI, United Microelectronics Corporation (UMC), Advanced Semiconductor Engineering Inc. (ASE), Compal Electronics (COMPAL), Taiwan Semiconductor Manufacturing (TSMC), AU Optronics (AUO) and so on are more favored by institutional and individual investors. Hence, the experiment selects 121 stocks. The processes are summarized as follows.

In XCS system, the condition part of the classifiers is considered into 4 parts in accordance with each of the discrete and normalized attributes. The detector will encode the inputted information into 12-bit strings. All classifiers will be matched. If the condition matches the current state, the rule will be activated, and then put into classifier list.

In XCS, the bucket brigade algorithm (Oh, Kim, & Min, 2005) is implemented for credit apportionment. In apportionment of credit system, matching classifiers are available for bidding. The fitness of the classifiers will be reflected through adjusting the strength. In Eq. (1), $B_{i,t}$ is the bid of classifier $X_i$ at time $t$, where $r$ is the learning rate, $S_i$ is the strength of the classifier, and $u_{i,t}$ is the number of non-wildcard symbols

$$B_{i,t} = \gamma \cdot S_i \cdot u_{i,t} \tag{1}$$

The probability that a classifier wins will be proportional to its bid value is shown in Eq. (2)

$$P(X_i wins) = \frac{B_{i,t}}{\sum B_{j,t}} \tag{2}$$

*4.3. Simulation result*

In Fig. 4, the profit of using XCS to allocate the MSCI component stocks during the training interval is remarkably increasing.

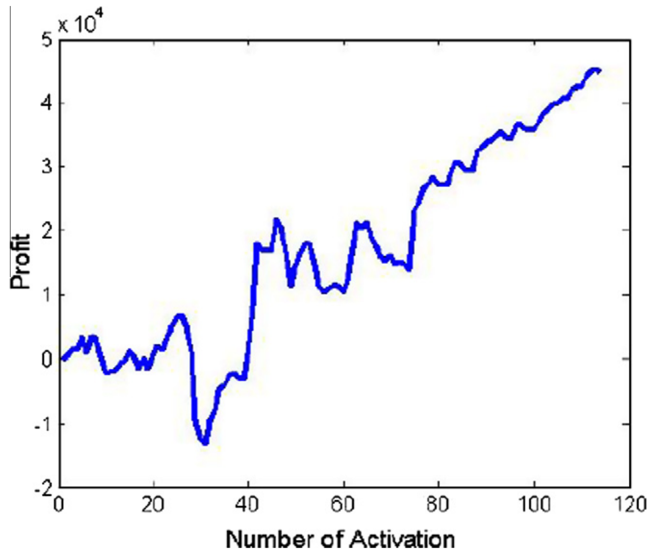


**Fig. 4.** Training period, the result is not good in the initial period, but the final profit can be achieved to 50 thousand dollars or so.

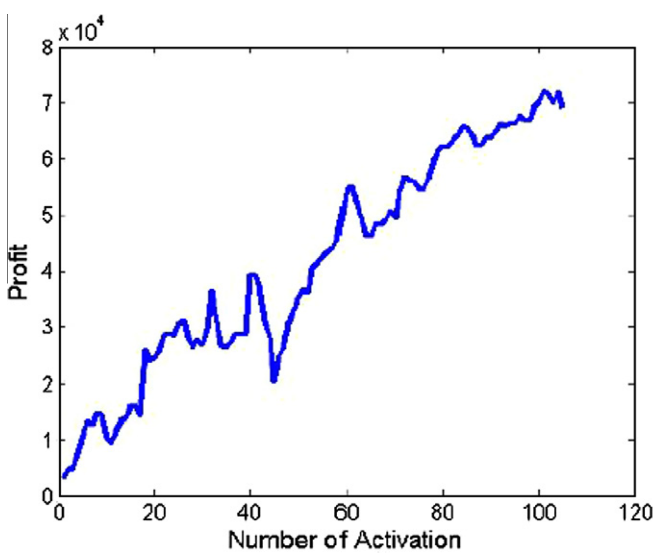


**Fig. 5.** Simulation period, the profit can be achieved to 71.3 thousand dollars or so.

In Fig. 5, the profit of using XCS to allocate the MSCI component stocks during the experimental report shows that we will gain a lot in the noisy stock market using XCS.

## 5. Conclusion

According to the simulation, the proposed model is the combination of XCS and stock market data to demonstrate that XCS can learn well from the complex, dynamic and noisy environment. XCS model possesses a strong advantage over traditional statistical analysis. The XCS provides guidance to stock market traders about how to adjust the weight of the MSCI component stocks. The profit according to the simulation is also satisfactory. Since the stock market is sensitive, using artificial intelligence is more appropriate because of continuous learning from the environment. So far, XCS works so adaptively that it is a good tool for further research of fundamental, technical, industrial, and news analysis on stock market. The simulation result shows that the accumulated return can be much higher than the original investment capital. Therefore, applying the mechanism proposed by this paper to stock data seems to be capable of profiting.

The profit gained through applying XCS to the financial markets is remarkable, because XCS is capable of adapting to the complex environment. Some works are still in the process of implementation for obtaining much better profit. First of all, the method of credit apportionment will be the focus of future researches. Recurrent reinforcement learning (RRL) and direct reinforcement (DR) will be replaced by bucket brigade algorithm to see whether the profit will be better. If the results are as much as desirable, the bucket brigade algorithm will be a brand new idea of implementing RRL or DR into classifier systems. In addition, Q-learning will also be compared with bucket brigade algorithm. Secondly, anticipatory classifier system (ACS) will also be applicable to continuous simulation (Stolzmann, 2000). The comparison of using different classifier systems will be discussed in future researches as well.

Last but not least, other attributes, such as ratio of margin purchasing and short selling (Oh et al., 2005; Trippi & Desieno, 1992), buy/sell of security dealers, TSE stock index, and volume of TAIEX, will be also considered to check whether the factors are affecting the financial market significantly.

## References

Chen, A.-P., & Chen, M.-Y. (2006). Integrating extended classifier system and knowledge extraction model for financial investment prediction: An empirical study. *Expert Systems with Applications, 31*, 174–183.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Reading, MA: Addison-Wesley.

Hashemi, R. R., Blanc, L. A., Rucks, C. T., & Rajaratnam, A. (1998). A hybrid intelligent system for predicting bank holding structures. *European Journal of Operational Research, 109*, 390–402.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor, MI: University Press of Michigan.

Holland, J. H., & Reitman, J. S. (1978). Cognitive systems based on adaptive algorithms. In D. A. Waterman & F. Hayes-Roth (Eds.), *Pattern directed interference systems* (pp. 313–329). New York: Academic Press.

Holmes, J. H. (1996). *Evolution-assisted discovery of sentinel features in epidemiologic surveillance*. PhD thesis. Philadelphia, PA: Drexel University.

Kovalerchuk, B., & Vityaev, E. (2000). *Data mining in finance*. Dordrecht: Kluwer.

Liao, P. Y., & Chen, J. S. (2001). Dynamic trading strategy learning model using learning classifier systems. In *Proceedings of the 2001 congress on evolutionary computation* (pp. 783–789).

McIvor, R. T., McCloskey, A. G., Humphreys, P. K., & Maguire, L. P. (2004). Using a fuzzy approach to support financial analysis in the corporate acquisition process. *Expert Systems with Applications, 27*, 533–547.

Oh, K. J., Kim, T. Y., & Min, S. (2005). Using genetic algorithm to support portfolio optimization for index fund management. *Expert Systems with Applications, 28*, 371–379.

Schulenburg, S., & Ross, P. (2001). Explorations in LCS models of stock trading. *Lecture notes in artificial intelligence*. Springer-Verlag (pp. 151–180).

Stolzmann, W. (2000). An introduction to anticipatory classifier systems. *Lecture Notes in Artificial Intelligence, 1813*, 175–194.

Studley, M., & Bull, L. (2007). Using the XCS classifier system for multi-objective reinforcement learning problems. *Artificial Life, 13*(1), 69–86.

Trippi, R. R., & Desieno, D. (1992). Trading equity index futures with a neural network. *Journal of Portfolio Management*(Fall), 27–33.

Tsai, W.-C., & Chen, A.-P. (2008). Service oriented architecture for financial customer relationship management. In *Proceedings of distributed event-based systems* (p. 332). ACM.

Wang, Y. F. (2003). Mining stock price using fuzzy rough set system. *Expert Systems with Applications, 24*, 13–23.

Wilson, S. W. (1996). Rule strength based on accuracy. *Evolutionary Computation, 3*(2), 143–175.

Yuan, Y., & Shaw, M. J. (1995). Induction of fuzzy decision trees. *Fuzzy Sets and Systems, 69*, 125–139.

Yuan, Y., & Zhuang, H. (1996). A genetic algorithm for generating fuzzy classification rules. *Fuzzy Sets and Systems, 84*, 1–19.