Automatic Trading System based on Genetic Algorithm and Technical Analysis for Stock Index

[1] Monruthai Radeerom
[1]*Faculty of Information Technology, Rangsit University,*
Pathumtani, Thailand 12000,
m.radeerom@rsu.ac.th

Abstract

Recent studies in financial markets suggest that technical analysis can be a very useful tool in predicting the trend. Trading systems are widely used for market assessment. This paper employs a genetic algorithm to evolve an optimized stock market trading system. Our proposed system can decide a trading strategy for each day and produce a high profit for each stock. Our decision-making model is used to capture the knowledge in technical indicators for making decisions such as buy, hold and sell. The system consists of two stages: elimination of unacceptable stocks and stock trading construction. The proposed expert system is validated by using the data of 15 stocks that publicly traded in the Thai Stock Exchange 100 Index (SET 100) from the year 2011 through 2014. The experimental results have shown Annual shape Ratio and Return Profits higher than "Buy & Hold" models for each stock index, and the models that used a genetic algorithm to selecting a trading signal has profit better than another models. The results are very encouraging and can be implemented in a Decision-Trading System during the trading day.

Keywords: *Computational intelligence, Genetic Algorithms, Stock Index, Technical Analysis*

1. Introduction

In recent financial markets, the use of automatic trading methods, which are often referred to as algorithmic trading, is expanding rapidly. Many works are found in applications of computational intelligence methodologies in finance [1]. Evolutionary computation, such as genetic algorithm (GA) [2], is promising in these methodologies, because of their robustness, flexibility and powerful ability to search.

Various works have been done on automated trading using evolutionary computation (e.g. [3 - 6]). These methods are mainly based on technical analysis, which is one of the two basic approaches in trading methods. Technical analysis is an attempt for the forecast of the future direction of prices by analyzing past market data, such as price and volume.

There is a large body of GA work in the computer science and engineering fields, but little work has been done concerning business related areas. Latterly, there has been a growing interest in GA use in financial economics, but so far there has been little research concerning automated trading. According to Allen and Karjalainen [7], genetic algorithm is an appropriate method to discover technical trading rules. Fern´ andez-Rodr´ıguez et al. [8] by adopting genetic algorithms optimization in a simple trading rule provides evidence for successful use of GAs from the Madrid Stock Ex-change. Some other interesting studies are those by Mahfoud and Mani [9] that presented a new genetic-algorithm-based system and applied it to the task of predicting the future performances of individual stocks; by Neely et al. [10] and by Oussaidene et al. [11] that applied genetic programming to foreign exchange forecasting and reported some success.

The aim of this study is to show how genetic algorithms, a class of algorithms in evolutionary computation, can be employed to improve the performance and the efficiency of computerized trading systems. It is not the purpose here to provide the theoretical or empirical justification for the technical analysis. We demonstrate our approach in a particular forecasting task based on emerging stock markets.

The paper is organized as follows: Section 2 presents the background about the Technical Analysis and genetic programming. Section 3 presents the GA decision-making model (GASignalTrade); Section 4 is devoted to experimental investigations and the evaluation of the decision-making model, and models the structure. The main conclusions of the work are presented in Section 5, with remarks on future.

2. Theoretical Background

2.1. Technical Analysis

Our approach to analyze financial markets is the technical analysis based on the past changes of prices and volume. The technical analysis needs various indicators for trading. They are evaluated from past stock data. Generally, technical indicators have several parameters. For example, moving average has a parameter, namely period, which is used as the denominator of averaging calculation. Various derived indicators, such as 10-days moving average, 50-days moving average, etc., are de-fined with the parameter.

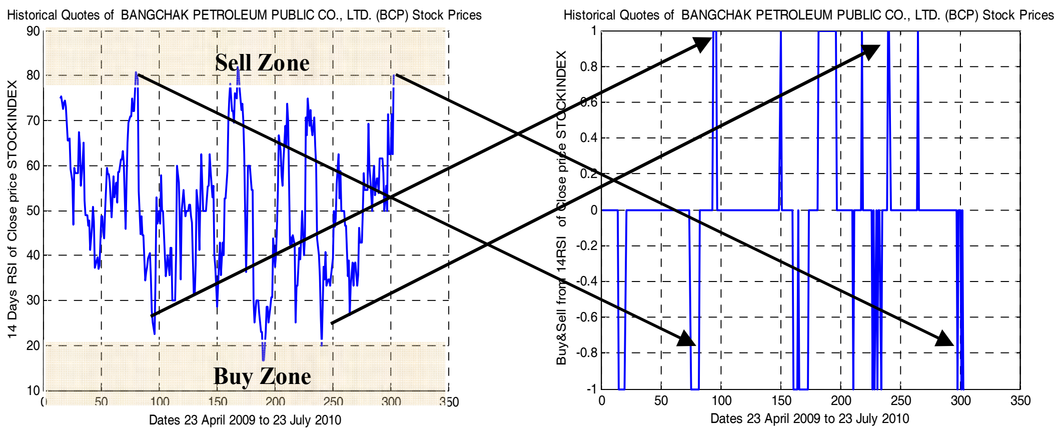
Relative Strength Index (RSI)

The Relative Strength Index (RSI) is a momentum oscillator used to compare the magnitude of a stock's recent gains to the magnitude of its recent losses, in order to determine the overbought or oversold conditions. The calculation formula

$$RSI = 100 - \frac{100}{1 + \frac{\sum(positive\ change)}{\sum(negative\ change)}},$$

Where RS=Average gains/Average losses.

Figure 1(a) shows a RSI index in 14 period's time. After that, we can evaluate buy, hold and sell signals by experience trader. Sell zone was RSI more than 70 and buy zone was RSI below 30. Thus, hold zone was between 30 and 70. Sell, buy and hold zone shown Figure 1(b).



(a) (b)

Figure 1. (a) The 14 periods RSI index (RSI14 (t)) calculated by close price(t)
(b) The Buy (1), Hold (0), Sell (-1) evaluated by RSI14(t)

Moving Average Convergence Divergence (MACD)

The MACD indicator constitutes one of the most reliable indicators within the market. It is a trend following the momentum indicator that exhibits the relation be-tween two distinct moving averages. Essentially, it defines two lines: the MACD line that corresponds to the difference between a 26-week and 12-week EMA and a trigger line that corresponds to an EMA of the MACD line. The difference between the former lines allows us to obtain a histogram that can be easily analyzed and offers us perspectives on price evolution.

$$MACD(S, L) = EMA(S) - EMA(l),$$

Where S = Lead Period's time, l = Lag Period's time, such as Lead 12 and lag 26 is S = 12, L = 26 respectively.

2.2. Risk-Insensitive Trader

For our trading system, the most natural objective function for a risk-insensitive trader is return profit and Annual Shape Ration. That is, the value obtained on the last investigation day is considered the profit. The trader's profit is calculated as

$$\text{Profit(n)} = \text{Stock Value(n)} - \text{Investment value},$$

Where n is the number of trading days.
And the Rate of Return Profit (RoRP) is

$$\text{RoRP} = \frac{\text{Profit(n)}}{\text{Investment value}} \times 100,$$

In finance, the Sharpe ratio (also known as the Sharpe index, the Sharpe measure, and the reward-to-variability ratio) is a way to examine the performance of an investment by adjusting for its risk. The ratio measures the excess return (or risk premium) per unit of deviation in an investment asset or a trading strategy, typically referred to as risk (and is a deviation risk measure). The Sharpe ratio is easily calculated using yearly data, or other time periods such as monthly or daily information. The higher a portfolio's Sharpe ratio, the more beneficial its returns have been historically, compared to its degree of investment risk. A ratio of 1 or better is considered good, 2 and better is very good, and 3 and better is considered excellent. The Sharpe Ratio is

Sharpe ratio = (investment's return % - risk-free return %) ÷ investment's standard deviation.

2.3. Genetic Programming

Genetic Programming is a branch of genetic algorithms. The difference between them is the way of representing the solution. Genetic programming creates computer programs as the solution whereas genetic algorithms create a string of numbers that represent the solution. Here the one-dimensional vector is called the chromosome and the element in it is a gene. The pool of chromosomes is called the population.
Genetic Programming uses these steps to solve problems.

i. Generate a population of random polynomials.
ii. Compute the fitness value of each polynomial in the population based on how well it can solve the problem.
iii. Sort each polynomial based on its fitness value and select the better one.
iv. Apply reproduction to create new children.
v. Generate new population with new children and current population.
vi. Repeat step ii – vi until the system does not improve anymore.

The final result that we obtain will be the best program generated during the search. We have discussed how these steps are implemented in our work in the next subsections.

Initial Population Generation: The initial population is made of randomly generate programs. We have used the traditional grow method of tree construction to construct the initial population. A node can be a terminal (value) or function (set of functions +, -, x, /, exp) or variable. If a node is a terminal, a random value is generated. If a node is a function, then that node has its own children. This is how a tree grows.

Fitness Evaluation: After the initial random population is generated, individuals need to be assessed for their fitness. This is a problem specific issue that has to answer "how good or bad is this individual?" In our case, fitness is computed by $\sum_{k=1}^{l}(g(p_k) - f_k)^2$ where k is the day in past, p is the past data, f is the present data and l is the length of the section found by concordance measures.

Crossover and Mutation: In a crossover, two solutions are combined to generate two new off spring. Parents for the crossover are selected from the population based on the fitness of the solutions. Mutation is a unary operator aimed to generate diversity in a population and is done by applying random modifications. A randomly chosen subtree is replaced by a randomly generated subtree. First, a random node is chosen in the tree, and then the node as well as the subtree below it is replaced by a new randomly generated subtree.

3. Methodology for The Intelligent Decision Trading system

Many stock market traders use conventional statistical techniques for decision-making in purchasing and selling [7]. Popular techniques use fundamental and technical analysis. They are more than capable of creating net profits within the stock market, but they require a lot of knowledge and experience. Because stock markets are affected by many highly interrelated economic, political and even psychological factors, and these factors interact with each other in a very complex manner. Figure. 2 show historical quotes of Bangchak Petroleum Public Co., Ltd. (BCP) stock prices. It is a high nonlinear system. In this paper, we are working on one-day decision making for buying/selling stocks. For that we are developing a decision-making model, besides the application of an intelligence system.

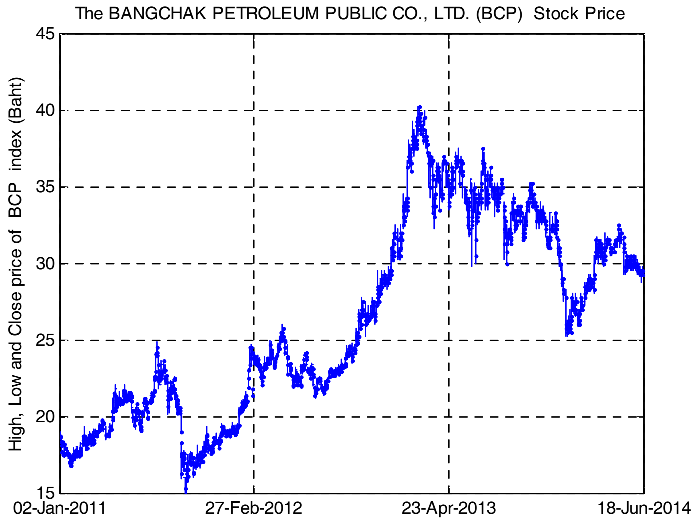


Figure 2. Historical Quotes of Bangchak Petroleum public Co., Ltd. (BCP) Stock Prices

3.1. Technical Indicators (TI)

Technical indicators are used by financial technical analysts to predict market movements, and many analysts use the standard parameters of these indicators. However, the best set of values for the parameters may vary greatly for each Thai Stock Index. Therefore, our approach uses GA to optimize these parameters considering the past time series history in order to maximize the profit. In this paper, we use five popular technical indicators: (i) Exponential Moving Average (EMA), (ii Moving Average Convergence/Divergence (MACD), (iii) Relative Strength Index (RSI) and (iv) Williams's %R (W). All of them correspond on open, high, low and close price.

3.2. Architecture and Trading Strategy

To analyze the proposed method, fifteen trading systems (TS) were created using the four technical indicators listed in 3.1, as shown in Table I. All the TSs considered are listed in Table I. The eleven remaining TSs were defined as a combination of four technical indexes. Each TS generated a trading signal for next day. For example,

TS1 : S = f(EMA),

Where S is trading signal (Sell(-1), hold (0), buy(1)).

From above, f(EMA) mean trading signal from evaluating by EMA based on maximum annual shape ratio. From maximum annual shape ratio, Lead and Lag which are parameters of EMA were adjusted.

TS6 : S = f(EMA+RSI),
Where S is trading signal (Sell(-1), hold (0), buy(1)).

From Equation above, f(EMA+RSI) mean average trading signal from EMA and RSI. Trading signal from EMA or RSI remains based on maximum shape ratio.

Many technical indicators have been proposed, but it is hard to select optimal signal trading from any indicators for actual algorithmic trading. Furthermore, it is also difficult to determine parameters for the selected indicators. In this paper, GAs is applied for this problem such as TS14.

TS14 : SGA = f(MACD+EMA+RSI),
Where SGA is trading signal (Sell(-1), hold (0), buy(1)).

Table 1. Input and output of Decision Trading System

TS	TS Rule	TS	TS Rule	TS	TS Rule
TS1	**S = f(EMA)**	**TS6**	**S = f(EMA+RSI)**	**TS11**	**S = f(EMA+RSI+W)**
TS2	**S = f(MACD)**	**TS7**	**S = f(EMA+W)**	**TS12**	**S = f(MACD+RSI+W)**
TS3	**S = f(RSI)**	**TS8**	**S = f(MACD+RSI)**	**TS13**	**S = f(MACD+EMA+RSI+W)**
TS4	**S = f(William)**	**TS9**	**S = f(MACD+W)**	**TS14**	**SGA = GA(f(MCAD+EMA+RSI))**
TS5	**S = f(EMA+MACD)**	**TS10**	**S = f(RSI+W)**	**TS15**	**SGA = GA(f(MACD+EMA+RSI+W))**

In the GA search, a genetic algorithm operate on a population of candidate trading signal encoded(S(MACD),S(EMA),S(RSI)). Each decision variable in the parameter set is encoded as a binary string and all are concatenated to from a chromosome as shown Figure 3. The precision of binary representation is tenth bits. It begins with a randomly constructed population of initial guesses. These solution candidates are evaluated in terms of our objective function. The objective function is used to measure risk-assessment of trading system. We selected Annual Shape Ratio. In order case, the GAs generated trading signal to maximize Annual Shape Ratio.

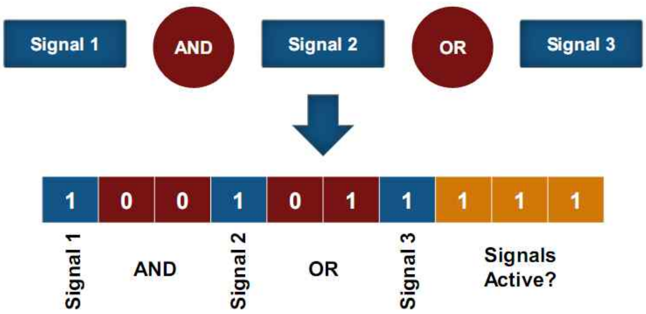


Figure 3. Combination of trading signal as bit string

The flow of our trading method is shown in Figure. 4. First, we apply the GA using stock price data for a predetermined period. This is the training phase to extract a set of effective technical indicators and their parameters. Next, we try an automated trading with the obtained set of technical indicators and their combined parameters using another stock price dataset. This is the testing phase to examine the performance of the selected indicators and their parameters. This process was executed with the 15 trading systems. Note that we do not apply the GA in the testing phase. It is important to prevent the over fitting in the training phase for improvement of the performance in the test.

The method of Figure 3 was implemented using the genetic algorithm of the MATLAB Global Optimization Toolbox [11], which has an implementation to handle Mixed Integer Optimization Problems [12]. This is used in the experiments to cope with the chromosome integer restrictions. The best profit in 50 executions of the training phase was chosen as the best fitted solution and then validated against the test data. The algorithm stops if there is no improvement in the objective function after at least 50 generations. The maximum number of generations chosen was 100 and two elite children are guaranteed to survive to the next generation.

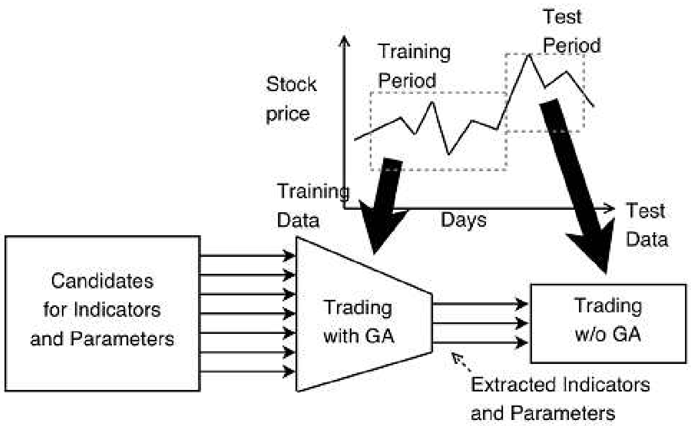


Figure 4. Flow of trading using GA.

During the training phase, the GA optimizes the technical indicator’s values of each trading system, aiming to maximize the Annual Shape Ratio by applying the signal trading from technical analysis defined by TS14 and TS15. The rules generate buy and sell signals. When no trading signal is emitted, the TS does not trade, staying in a neutral (standby) state. The trading systems were built using the stop-and-reverse strategy, that is, while the system is emit-ting a buy signal, the order is maintained intact until a sell or neutral signal is emitted, closing the current order and reverting the operation in case of a sell signal or just staying neutral. The same process occurs in the case of a recurrent sell signal output, reverting the order in the case of a buy signal or just closing the trade and staying neutral.

4. Results and Discussion

4.1. Setup

The model realization could be run having different groups of stocks (like Banking group, Energy group, etc.), indexes or other groups of securities. For that we are using market orders, as it allows simulating buying stocks when the stock exchange is nearly closed. All the experimental investigations were run according to the above presented scenario and were focused on the estimation of Rate of Return Profit (RoRP) and Annual Shape Ratio. At the beginning of each realization, the starting investment is assumed to be 1,000,000 Baht (Approximately USD 29,412). The data set, including 15 approved stock indexes in the Stock Exchange of Thailand (SET) index, has been divided into two different sets: the training data and test data. The stock index data is from January 4, 2011 to June 14, 2014 totaling 838 records. The first 720 records are training data, and the rest of the data, i.e., 118 records, will be test data. Moreover, the data for stock prices includes the buy-sell strategy, closing price and its technical data.

4.2. Stage 1: Elimination of unacceptable stocks

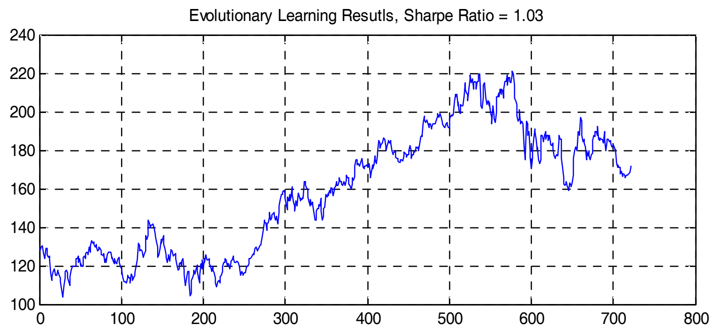
In this stage, the stocks that are not preferred by investors are eliminated. The unacceptable stocks are those that have a negative price to earnings ratio (P/E) or a negative shareholder’s equity value. For this reason, investors generally do not prefer to invest in these stocks. This stage reduces the burden on the stock evaluation stage, and prevents the system from suggesting unacceptable stocks to user. Finally, the proposed expert system is validated by using the data of 15 stocks that publicly traded in the Thai Stock Exchange 100 Index (SET 100) such as ADVANC, AEONTS, AOT, BAY, BCP, DTAC, JAS, KBANK, KKP, MINT, SCB, SPALI, TISCO, TNH, and TISCO.

4.3. Stage 2: Stock trading construction and Results

After developing the automatic trading system, we were given 1,000,000 baht for investment at the beginning of the training and testing period. The decision to buy and sell stocks is given by the proposed output (TS). Table 2 shows a Financial Simulation Model for calculating profit in our trading strategy. We calculated the Annual Shape Ratio and Return Profit that verify the effectiveness of the trading system. In Figure 5 and 6, TS14 including GA applied to BCP index. Parameters of this model is MA3, Lead 10, Lag 25, RSI 14 (30-70) and WR 46, respectively. Annual Shape Ration of Training set was 1.03 close to 1. And, return profit for 720 trading day was 119 %. In testing set, Annual Shape Ration of Training set was .885 close to 1. And, return profit for 118 trading day was 11.3 %.

Table 1. Example of Financial Simulation Model in trading strategy

GA14 : Buy & Sell (BCP)				Stock	Buy & Hold : Buy & Sell (BCP)			
Action		# of Shared	Cash(Baht)	Index	Action		# of Shared	Cash(Baht)
1	STAY	-	400,000	69.50	1	STAY	-	300,000
0	Hold	-	400,000	69.25	1	Buy	4,332	-
0	Hold	-	400,000	68.00	0	HOLD	4,332	
0	Hold	-	400,000	67.25	0	HOLD	4,332	
0	Hold	-	400,000	67.00	0	HOLD	4,332	
1	Buy	5,970	-	67.00	0	HOLD	4,332	
0	Hold	5,970	-	67.25	0	HOLD	4,332	
0	Hold	5,970	-	68.50	0	Hold	4,332	-
0	Hold	5,970	-	69.50	0	Hold	4,332	-
0	Hold	5,970	-	75.50	0	Hold	4,332	-
0	Hold	5,970	-	77.25	0	Hold	4,332	-
0	Hold	5,970	-	77.75	0	Hold	4,332	-
0	Hold	5,970	-	75.50	0	Hold	4,332	-



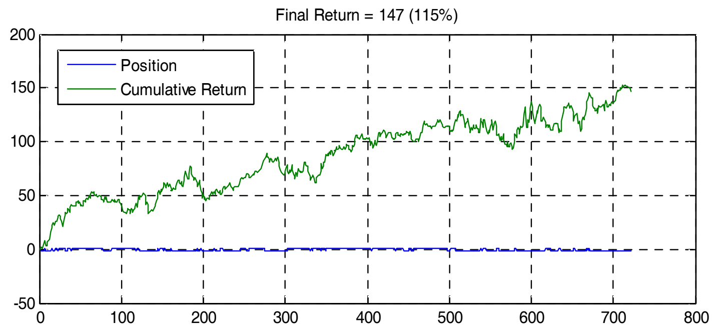


Figure 5. Annual Shape Ration and Return Profit of Training Set

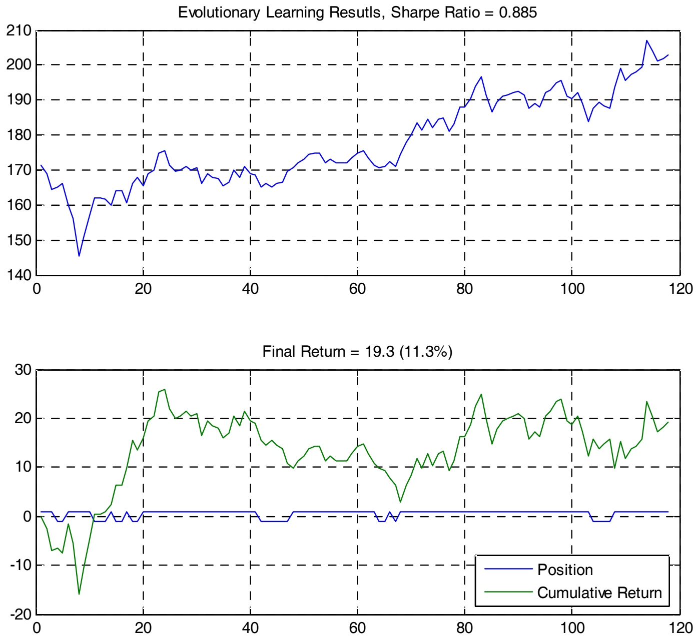


Figure 6. Annual Shape Ration and Return Profit of Test Set

Table 3. Summary of Annual Shape Ratio, Return Profit

Algorithm	Index	Average Profit				Buy & Hold			
		Annual Shape Ratio		Final Return		Annual Shape Ratio		Final Return	
		Train	Test	Train	Test	Train	Test	Train	Test
		771	68	771	68	771	68	771	68
TS1 : Lead/Lag EMA		0.815	-0.484	166.167	-2.649	0.235	-0.783	80.123	-3.154
TS6 : MA+EMA+RSI		1.009	-0.585	185.520	-3.057	0.345	-0.633	95.456	-3.652
S = (S1+S2)/2									
TS14 : MA+EMA+RSI		0.904	0.101	169.987	1.888	0.663	-0.323	101.236	-1.657
S = GAs(best,signals);									
TS15 : MA+EMA+RSI+WPR		1.200	0.306	208.533	3.485	0.733	-0.221	122.124	1.789
S = GAs(best,signals);									

Table 3 shown average Annual Shape Ratio and average Return profit from 15 stock indexes. TS14 and TS 15 trading rule have better than Ts1 – Ts13 not including GA.

Our experimental results show that GASignalTrade can improve digital trading by quickly providing a set of near optimum solutions. Concerning the effect of different GA parameter configurations, we found that an increase in population size can improve performance of the system. The parameter of the crossover rate does not seriously affect the quality of the solution.

5. Conclusion

This paper presented the decision-making model based on the application of GA. The model was applied in order to make a one-step forward decision, considering historical data of daily stock returns. The experimental investigation has shown a scenario of Intelligence Trading System based on Technical Analysis and GA to make a trading strategy, achieving more stable results and higher profits. For future work, several issues could be considered. Other techniques, such as support vector machines, particle swarm algorithms, etc. can be applied for further comparisons. This method should also be used to combine different signals that capture market dynamics more effectively (say a bear, bull, or side-ways market), or to analyze other stock index groups, other stock exchanges or other industries for comparisons. Another way, we really should add the window to the parameter sweep, or back test over our historical data and identify which training/validation window is optimal.

6. References

[1] A. Brabazon and M. O'Neill, "An introduction to evolutionary computation in finance," IEEE Computational Intelligence Magazine, pp. 42-55, 2008.

[2] D. E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, 1989.

[3] D. de la fuente, A. Garrido, J. Laviada and A. Gomez, "Genetic algorithms to optimize the time to make stock market investment," In Proc. of Genetic and Evolutionary Computation Conference, pp. 1857-1858, 2006.

[4] A. Hirabayashi, C. Aranha and H. Iba, "Optimization of the trading rule in foreign exchange using genetic algorithms," In Proc. 2009 IASTED International Conference on Advances in Computer Science and Engineering, 2009.

[5] M. A. H. Dempster and C. M. Jones, "A real-time adaptive trading system using genetic programming," Quantitative Finance, vol. 1, pp. 397-413, 2001.

[6] Matsui and H. Sato, "A comparison of genotype representations to acquire stock trading strategy using genetic algorithms," In Proc. Adaptive and Intelligent Systems, pp. 129-134, 2009.

[7] F. Allen, R. Karjalainen, "Using genetic algorithms to find technical trading rules," Journal of Financial Economic, Vol. 51, pp. 245-271, 1999.

[8] F. Fern´ andez-Rodr´ ıguez, C. Gonz´ alez-Martel, S. Sosvilla-Rivero. Optimization of Technical Rules by Genetic Algorithms: Evidence from the Madrid Stock Market, Working Papers 2001-14, FEDEA, [Online]. Available: ftp://ftp.fedea.es/pub/Papers/2001/dt2001-14.pdf, 2001.

[9] S. Mahfoud, G. Mani, "Financial forecasting using genetic algorithms," Journal of Applied Artificial Intelligence, vol. 10, pp. 573-565, 1996.

[10] (2012) MATLAB Global Optimization Toolbox [Online]. Available: http//www.mathworks.com/products/global-optimization/

[11] (2012) MATLAB Mixed Integer Optimization Problems, [Online]. Available: http://www.mathworks.com/help/toolbox/gads/bs1cibj.html#bs1cihn.

[12] C. Neely, P. Weller, R. Ditmar, "Is technical analysis in the foreign exchange market profitable? A genetic programming approach," In Proc. Forecasting Financial Markets: Advances for Exchange Rates, Interest Rates and Asset Management, London, 1997.