

Integration of Genetic Algorithm and Particle Swarm Optimization for Investment Portfolio Optimization

R. J. Kuo^{1,*} and C. W. Hong²

¹ Department of Industrial Management, National Taiwan University of Science and Technology, Taipei, Taiwan

² Innolux Corporation, Tainan County, Taiwan

Received: 25 Mar. 2013, Revised: 27 Jul. 2013, Accepted: 29 Jul. 2013

Published online: 1 Nov. 2013

Abstract: In investment market, investors often pay attention to investment portfolio selection and asset allocation under market risk. Thus, this study presents a two-stage method of investment portfolio based on soft computing techniques. The first stage uses data envelopment analysis to select most profitable funds, while hybrid of genetic algorithm (GA) and particle swarm optimization (PSO) is proposed to conduct asset allocation in the second stage. The evaluation results show that Sharpe value of portfolio based on the proposed method is superior to those of portfolio based on GA, PSO and market index. The proposed method really can help investors robustly obtain gains.

Keywords: Decision support system, Investment portfolio, Genetic algorithm, Particle swarm optimization.

1 Introduction

In recent years, financial market and social economy are changing sharply, hence, investors pay more attention to investment issues and investment products become more diversified. Mutual funds have the characteristics that investors share risks, are professionally managed and can be redeemed in short term, so investors prefer to invest it. For example, in Taiwan, the number of equity funds is increasing from 246 in 2005 to 299 in 2008. Fund size is also increasing, which amount accounts for the half. Thereby, equity funds play an important role in mutual fund in Taiwan.

In discussion of portfolio, portfolio selection, asset allocation, and cash reserve ratio are very important and also make investors headachy. The traditional approaches to solve portfolio problems need application of many complex statistical methods and reference variables provided by many experts, so they are not adopted by the people. In addition, the traditional tools lack timeliness. For investment market, every second counts, and delay is not allowed. Investors stand in need of timely investment suggestion.

Recently, thanks to the rapid development of computer technologies, soft computing techniques are widely used to solve many complicated problems. More

and more investment consulting companies apply soft computing techniques to analyze data in company financial reports and industry information, and provide investment guidelines.

Therefore, this study intends to propose a two-stage method that is able to select the funds with better performance and determine the most suitable asset allocation. It can help investors who are with a little financial knowledge hold portfolio with low risk and high gains. This study uses data envelopment analysis (DEA) to estimate efficiency of equity funds in financial market, and then hybrid of genetic algorithm (GA) and particle swarm optimization (PSO) in asset allocation, so as to build highly robust portfolio.

The rest of this paper is organized as follows. Section 2 presents the backgrounds related to this study, while the proposed two-stage method is explained in Section 3. Section 4 illustrates the model evaluation results. Finally, the concluding remarks are made in Section 5.

2 Literature Survey

This section will briefly present the backgrounds necessary for the current study. It includes portfolio, mutual fund, DEA, GA, PSO, and hybrid of GA and PSO.

* Corresponding author e-mail: rjkuo@mail.ntust.edu.tw

2.1 Portfolio Selection Problem

Portfolio is the means of diversification of risk in financial market. Research on portfolio began as early as 1952 and proposed by Markowitz. Portfolio is a collection of more than one assets or securities. The portfolio focuses on diversification of risk, maximization of return rate on investment and evaluation of risk and performance. When investors are uncertain of investment decisions, they will consider reduction of investment opportunities to reduce risk or invest different ones for diversification of risk. As for investors, the focus of attention is not volatility of individual security but the total return rate and risk of the portfolio of securities. The modern portfolio theory can facilitate development of mutual funds and also make investors attain the high return and bear low risk.

In the modern portfolio theory, Markowitz firstly put forward the portfolio in Portfolio Selection in 1952, in which he suggested the concept of qualifying risk, and risk should be added in the investment decision. The return rate on the securities is calculated with average value, the risk is estimated by variance of return rate. Aside, the degree of risk correlation among securities in portfolio is measured by covariance. The most suitable portfolio model is the efficient set or efficient frontier, the conditions of portfolio which meet the efficient set are: return rate can be maximized when risk is fixed, or risk is minimized when return rate is fixed.

Based on Markowitz theory, Sharpe (1963) suggested Capital Asset Pricing Model (CAPM). The development of this theory includes two parts, one is capital market line, and the other is security market line. Both of them are derived from the concept of a market portfolio. The method assumes that risk is linearly related to return. The reward for holding a risky investment exceeding a risk-free one is called risk premium. Return rate on all securities is only affected by a single factor, and that is market portfolio.

Sharpe (1963) suggested investment of any risky security or asset (real estate) bears system risk and unsystematic risk. Portfolio can eliminate unsystematic risk which is also called diversifiable risk. Statman (1987) proved portfolio can really diversify risk.

2.2 Mutual Fund

Mutual fund is a professionally managed type of collective investment scheme that pools money from many investors and invests it in securities to establish a diversified portfolio of investments. This is an indirect investment, and investors participate proportionally in the gain or loss of the fund.

In terms of investment objects, funds can be divided into equity fund, balanced fund and bond fund. Equity fund will be the main focus for this study. Equity fund is the mutual fund which investment objects are listed stock and over-the-counter stock.

According to Regulations of Governing Securities Investment Trust Funds, equity fund refers to a fund in which total investment in stocks accounts for no less than 70 percent of the total asset value of the fund. When the name of an equity fund indicates investment in a particular investment vehicle, region, or market, then investment in the vehicle, region, or market so indicated shall account for no less than 60 percent of the total asset value of the fund. According to investment objective and scope, equity fund can be subdivided into several types, such as high-tech fund and index fund, etc. Equity fund can be divided into domestic investment and overseas investment.

Ippolito (1989) extended CAPM by Sharpe (1963). In addition to the original market portfolio factors, the fund turnover rate, management fees and other relevant expense rate are added, together with whether mutual fund receives sales charge and dummy variables of current year. In terms of results, fund turnover rate, management fees and other relevant expense rate and current year are not significantly related to the mutual fund performance. Smith and Tito (1969) studied 38 types of open-end funds in US between 1958 and 1967, and used performance indicators by Sharpe, Treynor, Jensen and the one adjusted by Jensen to measure performance of funds. The results reveal that mutual fund performance is better than the market portfolio performance. The above four performance indicators have high correlation. Fama (1972) suggested two sources of excess return of two types of mutual funds: selectivity and market timing ability. Grinblatt and Titman (1992) studied 279 types of mutual funds between 1974 and 1984 and used regression model to estimate excess return and slope coefficient of excess return. Based on empirical results, they illustrate that performance of mutual funds has positive correlation. Brown and Goestmann (1995) took 372 to 829 open-end funds in US in 1976-1988 as subjects and applied statistical validation, and found historical information can help analyze fund performance.

Wermers (2000) investigated the performance level of 2400 types of mutual funds in US between 1974 and 1994, and considered active investment management of fund manager is not necessary according to cost-benefit ratio. Droms and Walker (2001) used the equity fund data from US between 1971 and 1990 to integrate time sequence and cross-section model for analysis of performance of equity funds, and found the performance persistence exists in three years and is not significant beyond four years. Annaert et al. (2003) took European equity funds over the period 1995-1998 as subjects and adopted Bayesian stochastic frontier approach for analysis. The empirical results show those European equity funds have economies of scale, and efficiency is related to historical performance. Santos et al. (2005) employed stochastic frontier approach to evaluate 307 Brazilian equity funds over the period 2001-2003. The results indicate the fewer changes the portfolio has, the

higher efficiency is. Basically, the performance of funds is unrelated to scale.

Different measurement indexes lead to different performance evaluation model of mutual funds. Generally, the method of evaluating performance of mutual funds is mainly based upon CAPM, that is to say risk-adjusted return rate on investment serves as index to measure fund performance.

The traditional evaluation indicators are mainly (1) Treynor performance indicator, (2) Sharpe performance indicator and (3) Jensen performance indicator, and the above three indexes are derived from CAPM.

1. Treynor performance indicator

Treynor (1965) considered that system risk can be used to measure performance of portfolio. Thus in terms of security market line, system risk stands for risk coefficient to adjust return on portfolio and further to derive Treynor performance indicator.

$$T_i = \frac{E(R_i) - E(R_f)}{\beta_i} \quad (1)$$

T_i : Treynor performance indicator for mutual fund portfolio.

$E(R_f)$: risk-free interest in market.

$E(R_i)$: average value of portfolio.

β_i : system risk value.

Treynor performance indicator is the extra return obtained by investor when bearing a unit of system risk. The greater T value is, the larger the extra return obtained by investor when bearing a unit of system risk. The system risk value β stands for volatility of portfolio, the greater the β is, the greater volatility of portfolio is with volatility of market portfolio.

2. Sharpe performance indicator

Sharpe (1966) suggested relation between risk and return should be considered in evaluation of portfolio performance. However, Treynor performance indicator only considers system risk and ignores nonsystematic risk. Thus Sharpe used standard deviation (total risks, including system risk and nonsystematic risk) of return rate on portfolio as risk coefficient to adjust return rate on portfolio, which serves as measurement method of portfolio. The Equation is as follows:

$$S_i = \frac{E(R_i) - E(R_f)}{\delta_i} \quad (2)$$

S_i : Sharpe performance indicator for mutual fund portfolio.

δ_i : stands for standard deviation of return rate on the i of portfolio.

$E(R_i)$: average value of portfolio.

$E(R_f)$: risk-free interest in market.

3. Jensen performance indicator

Jensen (1968) adopted concept of absolute performance indicator, and the performance evaluation of mutual fund portfolio and of market portfolio under the same risk level represents difference of return rate on mutual fund portfolio and market portfolio under the same risk level, mainly evaluating whether mutual fund portfolio is higher than market portfolio under the same risk level. Thus J may be negative or positive. The Equation is as follows:

$$R_i - R_f = J_i + (R_m - R_f) \beta_i \quad (3)$$

J_i : Jensen performance indicator for mutual fund portfolio.

R_m : average market return rate.

2.3 Data Envelopment Analysis

Data envelopment analysis (DEA) developed by Charnes, Cooper and Rhodes (Banker et al., 1984 and Charnes et al., 1978) is a mathematical programming technique for measuring the relative performance of decision making units (DMUs) on the basis of the observed operation practice in a sample of comparable DMUs. It has been widely applied to analyze the relative production efficiency of DMUs in a setting of multiple incommensurate input and output variables. A lot of researchers used DEA in many application since DEA was presented in 1978. In the following, applications of DEA to mutual funds will be discussed.

Basso and Funari (2001) suggested DEA-g model measured performance of 50 funds in Italian financial market, and g-model is multiple inputs and multiple outputs. The input variables are standard deviation and investment cost (subscription and redemption), and output variables are extra return, stochastic dominance indicator, Sharpes indicator, Treynors indicator and Jensens indicator. Computational results indicate DEA-g can take different views into account in fund performance. It is more suitable than only considering one single indicator, like Sharpes indicator, Treynors indicator and Jensens indicator, half-var, DEA-1 and DEA-2. Murthi et al. (1997) studied performance of individual funds with reference to 7 types of US funds totaling 731 in 1993 and 33 types of mutual funds totaling to 2083 in the third quarter. They define that input variables are trading fees, commission charge, turnover rate and standard deviation and output variable is excess return rate. This study makes comparison between DEA and Jensens indicator and Sharpes performance indicator. McMullen and Robert (1998) studied 135 US equity funds in 1997, and the input variables contain

standard deviation, minimum investment amount, turnover rate, and sales rate, while output variables include return rate of 1 year, 3 years and 5 years. The empirical results show large fund has less efficiency, because investment has no corresponding return.

2.4 Genetic Algorithm

Holland proposed genetic algorithm (GA) in 1975, and its basic view is based on survival of the fittest in natural selection- Charles Darwin. Based on the survival and reproduction of the fittest, GA adopts a group of simulated encoded chromosomes and calculates the fitness of chromosomes. Each chromosome undergoes crossover and mutation to produce next generation. This evolution process continues until the stopping criteria are reached. GA is appropriate for large-sized and nonlinear space problems whose solutions are unpredictable. Relying on multi-point search and algorithmic features, it is not easy to fall into local optimal solution but can converge to universal optimal solution. Kaboudan (2000) adopted GA to charge off intraday stock to forecast highest and lowest price for every day. The results indicate gain of forecast individual stocks is up to 20% 80% in 20 days. Oh et al. (2006) took KOSPI200 as object, and gave different weights to standard deviation, portfolio β , average trading volume according to market conditions to design a series of index funds. Kim and Han (2006) combined neural network with GA to simulate optimization of stock market. Maringer and Kellerer (2003) studied optimization of portfolio by using simulated annealing (SA) method and GA. Kuo et al. (2001) tried to combine GA-based fuzzy neural network and neural network to develop a stock trading decision support system. The computational results revealed that GA-based fuzzy neural network is able to reflect the current economic situation. Chen et al. (1988) applied market trading data of S&P500 index options to genetic programming method. The results show genetic programming has good forecast performance.

2.5 Particle Swarm Optimization

Particle swarm optimization (PSO) was put forward by Eberhart and Kennedy (1995). The concept is mainly from collective animal behavior (Boeringer and Werner, 2004). Although its development is a little late compared to GA, it is now applied to solution of optimization of many portfolios. Particle swarm contains two concepts, one is that Boyd and Richerson (1985) proposed individual will refer to their own experience or experience of others in decision making according to the human decision process. The other is to propose simple rules to

modularize collective natural behavior. Reynolds (1987) thought complicated collective behavior can be simulated by the three following aspects: follow the individual most closed to objects, move towards object, and move toward group center.

In the original PSO, particle is represented as $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$, which represents a potential solution to a problem in D-dimensional space. Each particle keeps a memory of its previous best position P_{best} , and a velocity along each dimension, represented as $v_i = (v_{i1}, v_{i2}, \dots, v_{iD})$. At each iteration, the position of the particle with the best fitness value in the search space, designated as g , and the P vector of the current particle are combined to adjust the velocity along each dimension, and that velocity is then used to compute a new position for the particle (Eberhart and Kennedy, 1995 and Wang et al., 2005). The method could be divided into GBEST and LBEST versions (Eberhart and Kennedy, 1995), whose main difference is their definition of the best. In the GBEST version, the particle swarm optimizer keeps track of the overall best value, and its location, obtaining thus far by any particle in the population, which is called $g_{best}(G_{bestid})$. For the LBEST version, in addition to g_{best} , each particle keeps track of the best solution, called $l_{best}(G_{bestid})$, and it is attained within a local topological neighborhood of particles. However, the particle velocities in each dimension are held to a maximum velocity, v_{max} , and the velocity in that dimension is limited to v_{max} . The updating rule is as follows:

$$v_{id}^{new} = v_{id}^{old} + c_1 \times rand_1 \times (P_{bestid} - X_{id}) + c_2 \times rand_2 \times (G_{bestid} - X_{id}) \quad (4)$$

$$X_{id}^{new} = X_{id}^{old} + v_{id}^{new} \quad (5)$$

where c_1 and c_2 determine the relative influence of the social and cognition components (learning factors), while r_1 and r_2 denote two random numbers uniformly distributed in the interval [0,1]. After the first version of PSO was proposed, many efforts have been made to improve the performance of PSO, like (Shi and Eberhart, 1998a; Shi and Eberhart, 1998b; Xie et al., 2002; Bergh and Engelbrecht, 2002; Bergh and Engelbrecht, 2003; Liu et al., 2007; Arumugam and Rao, 2008).

2.6 Hybrid of GA and PSO

Both GA and PSO have its own merits. In order to combine the advantages of both methods, there already have been many researches working on this area. Juang (2004) improved mathematical calculation effect of

recurrent neural network by combining GA with PSO. Shi et al. (2005) introduced GA into calculation process of PSO and calculated its adaptive value, and compared the value with single GA and PSO. The results indicate hybrid method has better optimization performance. Du et al. (2006) set GA operational factor into PSO position updated to get rid of possible local solution. The demonstration finds it can improve search and increase sophistication of solution. Laura and Mihai (2007) took fund portfolio problem as an example, and compared combination of GA and PSO with standard PSO. After verification, the combination has significant performance. Kao and Zahara (2007) applied a hybrid GA and PSO for multimodal function, in which initial matrix is divided into two parts. GA and PSO is carried out separately, and then crossover, and repeated solution.

3 Methodology

This study firstly collected the financial reports in listed companies in semiconductor industry and conducted DEA of the data in the reports with reference to financial indicators used by literatures and selected the components, funds, which are qualified for portfolio. Next, two hybrids of GA and PSO are proposed to allocate the portfolio. The research process is illustrated in Figure 1.

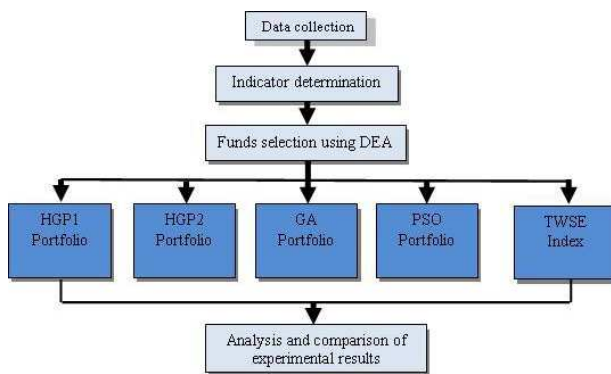


Fig. 1: Research flowchart.

3.1 Fund Selection Strategy

3.1.1 Data collection

In investment market, investors select funds on the basis of different performance indicators. This study takes equity funds as subjects and refers to performance variables by McMullen and Robert (1997) and

Generalized DEA by Basso and Funari (2001). The data are from <http://www.fin.ntu.edu.tw/>.

3.1.2 Performance indicators selection

This study determined the performance indicators by referring to fund performance indicators used in (McMullen et al., 1998) and (Basso and Funari, 2005). Since DEA is applied to select qualified funds for portfolio with adjusted fund performance variables, it is necessary to differentiate performance variables. The smaller the input variables and the larger the output variables, the better, while the output variables should be the larger the better. Then CCR (Charnes et al., 1978) model is employed to calculate the efficiency values of these companies according to their corresponding input and output variables. Finally, using the efficiency values to carry out fund sorting.

3.2 Model Construction

After qualified funds have been selected by using DEA, this study further utilizes the selected funds to constitute a financial model and construct portfolio with optimal performance. The training data and testing data are 11 months and 1 month, respectively. This study proposes hybrid GA and PSO to solve portfolio. We firstly construct basic portfolio return rate as follows:

$$R_t = \sum_{k=1}^n w_k E(r_k) \tag{6}$$

$$0 < w_k < 1 \tag{7}$$

$$\sum_{k=1}^n w_k = 1 \tag{8}$$

Use Sharp equation as fitness function which is as follows:

$$Max \quad S = \frac{M(R_t) - R_f}{\sqrt{v(R_t)}} \tag{9}$$

$$M(R_t) = \frac{1}{T} \sum_{t=1}^T R_t \tag{10}$$

$$v(R_t) = \frac{1}{T} \sum_{t=1}^T [R_t - M(R_T)]^2 \tag{11}$$

S : fitness function.

R_f : month fixed interest rate of fixed deposit in bank.

$M(R_t)$: monthly return rate of portfolio.

$v(R_t)$: standard deviation of portfolio

This study adopts buy-and-hold strategy, so the weight given must be larger than zero. To avoid total

weight w_k greater than 1 in calculation, weight is normalized when total weight is greater than 1. In other words, make each weight in $[0,1]$ and total value be 1. For the fitness function, this study focuses on expected return of fund and total variance of portfolio, and intends to gain high return under risk. The higher fitness function value is, the better performance is. Further, no limitation is made to weight of each fund with a view to selecting stock portfolio with better performance.

3.3 Hybrid GA and PSO to Solve Portfolio

The method for solving portfolio is hybrid of GA and PSO. Next we introduce two types of hybrid models. For the proposed hybrid methods, the floating-point encoding method based on Wang et al. (2005) is used. The main idea of the proposed methods is to combine mutation, which can help avoid stuck in local optimal solutions, and elitist strategy which is able to speed up the evolution with PSO according to (Kao and Zahara, 2007). This can provide better performance.

3.3.1 Hybrid GA and PSO- HGP1

The computational process for HGP1 is as follows:

Step 1: Set up algorithm parameters.

Step 2: Set up search range for X , which has impact on search speed. Initial speed X_{id} of particles is generated at random in given range. The initial speed of V_{id} is restricted to random generation in V_{max} .

Step 3: Calculate fitness function value of all particles.

Step 4: Record and update individual extreme value (optimal position and optimal fitness value of each particle) and extreme value of population (optimal position and fitness value of all particles).

Step 5: Sort particles according to fitness values (from high to low), and equally divide them into two halves.

Step 6: Use first half of particles to update the present particles and substitute the second half of particles, and the Equation is as follows:

$$v_{id}^{new} = \omega \cdot v_{id}^{old} + c_1 \cdot rand_2 \cdot (p_{id}^{old} - x_{id}^{old}) + c_2 \cdot rand_3 \cdot (p_{id}^{old} - x_{id}^{old}) \quad (12)$$

$$x_{id}^{new} = x_{id}^{old} + v_{id}^{new} \quad (13)$$

Step 7: Conduct crossover and mutation for all particles. Crossover is implemented based on (Kao and

Zahara, 2007) and the equations are as follows:

$$x_{id}^{new1} = uniform(0,1) \cdot x_{id} + (1 - uniform(0,1) \cdot x_{id+1}) \quad id = 1, 2, \dots, n-1 \quad (14)$$

$$x_{id}^{new1} = uniform(0,1) \cdot x_{id} + (1 - uniform(0,1) \cdot x_{id}) \quad id = n \quad (15)$$

The mutation equation is as follows:

$$x_{id}^{new2} = x_{id}^{new1} + rand_1 \times N(0,1) \quad (16)$$

X_{id} of each particles is limited to X range, and let X_j be between upper and lower limit:

$$0 \leq x_i \leq 1 \quad (17)$$

Step 8: calculate fitness function value of all particles. After comparison of the half of particles not to be updated with those after mutation, and select particles with better fitness value by using Elitist Strategy

Step 9: Stop if the specified iteration number is reached; otherwise, go back to Step 3.

3.3.2 Hybrid GA and PSO- HGP2

The computational process for HGP2 is as follows:

Step 1: Set up algorithm parameters.

Step 2: Set up search range for X , which has impact on search speed. Initial speed X_{id} of particles is generated at random in given range. The initial speed of V_{id} is restricted to random generation in V_{max} .

Step 3: Calculate fitness function value of all particles.

Step 4: Record and update individual extreme value (optimal position and optimal fitness value of each particle) and extreme value of population (optimal position and fitness value of all particles).

Step 5: Sort particles according to fitness values (from high to low), and equally divide them into two halves.

Step 6: Perform crossover and mutation of GA for first half of particles to update particles, and calculate fitness value of each particle to replace second half of particles. x_{id} of each particle is limited by X range, and let x_i be between upper and lower limit:

$$0 \leq x_i \leq 1 \quad (18)$$

Crossover is conducted using the equation proposed by Kao and Zahara (2007) as follows:

$$x_{id}^{new1} = uniform(0,1) \cdot x_{id} + (1 - uniform(0,1) \cdot x_{id+1}) \quad id = 1, 2, \dots, n-1 \quad (19)$$

$$x_{id}^{new1} = uniform(0, 1) \cdot x_{id} + (1 - uniform(0, 1) \cdot x_{id}) \cdot id = n \tag{20}$$

Mutation equation is as follows:

$$x_{id}^{new2} = x_{id}^{new1} + rand_1 \times N(0, 1) \tag{21}$$

Step 7: Use results in Step 3 to update first half of particles, and equation is as follows:

$$v_{id}^{new} = \omega \cdot v_{id}^{old} + c_1 \cdot rand_2 \cdot (p_{id}^{old} - x_{id}^{old}) + c_2 \cdot rand_3 \cdot (p_{id}^{old} - x_{id}^{old}) \tag{4}$$

$$X_{id}^{new} = X_{id}^{old} + v_{id}^{new} \tag{5}$$

Step 8: Stop if the specified iteration number is reached; otherwise, go back to Step 3.

4 Model Evaluation Results and Discussion

This section takes Taiwan equity funds as an example for the purpose of validation. Taiwan equity funds increase considerably. Currently, there are more than 100 types of funds. It is necessary to have a scientific way to help investors make right decision. In the first phase, we find out the financial indicators containing performance connotation based on the literature, and employs DEA to analyze these financial indicators in order to select suitable funds with better performance. In the second phase, we apply the two proposed methods, HGP1 and HGP2, GA and PSO to determine effective asset allocation according to Sharpe value.

4.1 DEA Variables

Funds have the characteristics of multiple input and output and results are closely related to whether suitable input and output variables are selected or not. With reference to generalized DEA presented by Basso and Funari(2001), input and output variables suggested by McMullen and Robert(1997), and Taiwan equity funds, we select four input variables including commission, transaction tax rate, monthly standard deviation and β value, and four output variables including Sharpe's, Jensen's, Treynor's ratio and excess return rate.

4.2 Data Source and Preprocess

The data in validation analysis is from monthly fund data provided from <http://www.fin.ntu.edu.tw/>. The data period is between February 2005 and December 2008. Some funds whose input or output data are not complete have been deleted. In all, the number of funds is about 140~160, which meets the principle that number of decision making units must be larger than two times of number of input and output variables.

4.3 DEA Results

After using DEA to analyze the decision making units with the corresponding input and output data, top 15 funds with good performance in each month are selected. Totally, there are 47 months for the current study. Each fund is ranked. Since some funds are renamed, we use the name of that period. The top 15 funds over the period of 01/2006~02/2006 are listed in Table 1.

4.4 Parameter Design

HGP1 and HGP2 in this study need parameter setup for assessment. HGP1 and HGP2 are hybrid of GA and PSO, so this study has to set up parameters including crossover rate, mutation rate, maximum speed and inertial weight. The optimal parameters are obtained by using Taguchi method.

The Taguchi experiment is a kind of partial factor experiment. Orthogonal array consists of factors and levels. For this current study, the parameters are four control factors, and meanwhile parameters can be designed in combination of three levels. Basically, the parameter level design will focus on four control factors and comply with the literature. They are summarized in Table 2.

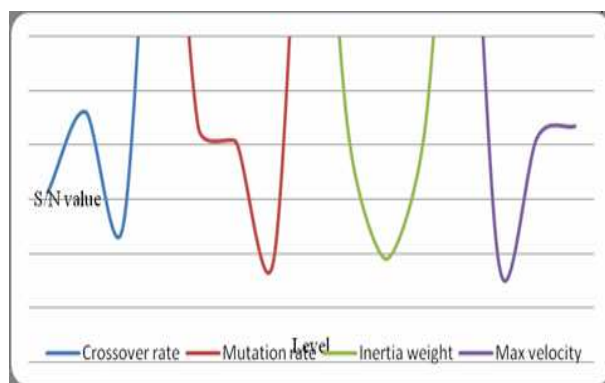
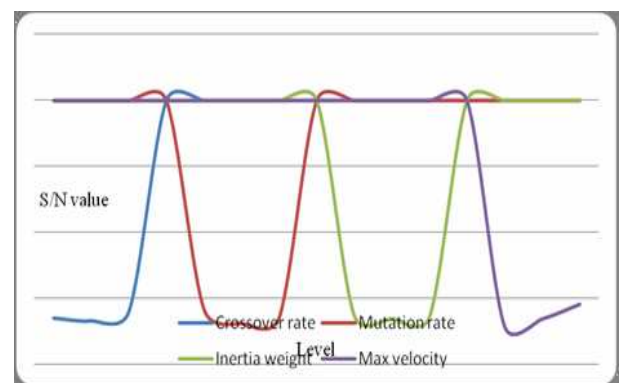
Table 2: Control factors and level values.

Crossover rate	1	0.6	0.8
Mutation rate	0.2	0.3	0.1
Inertia weight	0.5+rand/2	0.9~0.4	0.72
Maximum velocity	10	5	0.4

The data is fund data in January 2006, and the parameter combination with best fitness value is the parameters in this study. In consideration of practice and computational cost, Taguchi experiment is stopped after optimal solution is calculated fifty times and no change occurs for HGP1. The experiment is performed 10 times. The other parameters are number of particles: 40, C1 and C2 is 2 respectively, and S/N value is shown in Figure 2. The optimal parameters of HGP1 after ten-time experiment are as follows: crossover rate is 0.6, mutation rate is 0.2, maximum speed is 0.4, and inertial weight is 0.72. In addition, Taguchi experiment is performed 10 times for HGP2. On account of time, the experiment is stopped after optimal solution is calculated fifty times and no change occurs. And other variables are number of particles: 40, C1 and C2 is 2 respectively. The optimal parameters of HGP2 are as follows: crossover rate is 0.8, mutation rate is 0.2, and maximum speed is 0.4, and inertial weight is 0.5+rand/2, and S/N value is shown in Figure 3.

Table 1: Top 15 funds which performance value ranks first.

No.	Fund (01/2006)	Efficiency value	No.	Fund (02/2006)	Efficiency value
1	JF(Taiwan) Value Growth	1	1	Jih Sun Top Five	1
2	HSBC Taiwan Blue-Chips	1	2	JF(Taiwan) Value Growth	1
3	Cathay Cathay	1	3	SinoPac	1
4	FUHWA	1	4	ING Taiwan Aggressive Growth Selection Fund	1
5	INVESCO Taiwan Technology	1	5	Fuh-Hwa Fuh-Hwa	1
6	Jih Sun Neo Taiwan Enterprises	1	6	INVESCO Mainstream Equity	1
7	Fidelity Taiwan Growth	1	7	HSBC Taiwan Blue-Chips	1
8	Cathay Small & Medium Cap	1	8	HSBC Taiwan Mid & Small Cap	1
9	ING Taiwan Aggressive Growth Selection Fund	1	9	FUHWA	1
10	SinoPac	1	10	Prudential Financial OTC	1
11	Chung Hsing Small Capital	1	11	AIG Taiwan Giant	0.977
12	Jih Sun Top Five	1	12	Fidelity Taiwan Growth	0.956
13	Franklin Temp First Taiwan First	1	13	PCA ESSENCE	0.947
14	Prudential Financial OTC	1	14	TaiwanChina Focus Fund	0.917
15	AIG Taiwan Giant	0.989	15	Cathay Cathay	0.903

**Fig. 2:** Level S/N value response of HGP1 factors.**Fig. 3:** Level S/N value responses of HGP2 factors.

4.5 Asset Allocation Results

The experimental periods include years 2006-2008. There are 12 months in each year, and the experiment is performed thirty times per month. Five Sharpe values are used to compare portfolio performance. X axis and Y axis

are month and Sharpe value, respectively. Figure 4 lists results of the first period. It shows that the performance of the four kinds of portfolio is all better than market return. HGP1 is the best, followed by HGP2. For number of winning months, HGP1 has 12 months. For the average

Sharp value of the 12 months in the first period, HGP1 is also the first, HGP2 is second, GA third, PSO fourth. Table 3 presents all these experimental results.

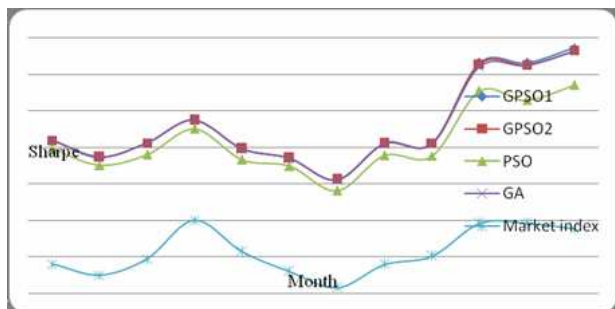


Fig. 4: Experimental results of first period (2006).

Table 3: Average Sharpe value and number of winning months in the first experiment (2006).

Portfolio	HGP1	HGP2	PSO	GA	Market Index
Average Sharpe value	0.920035	0.916789	0.830507	0.914638	0.22824
Number of winning months	12	0	0	0	0

In the second experimental period (2007), market changed sharply. The performance of 12 months based on four algorithms is compared with the market return, see Figure 5. The performance value of the four kinds of portfolio is better than the market return; for the number of winning months, HGP1 has 12 months. For the average monthly Sharpe value over the second research period, HGP1 is the first, HGP2 is second, GA third, PSO fourth. But for number of winning months, GA is superior to PSO, and PSO is superior to GA due to Sharpe value depends on interaction between excess return rate and risk, so return rate is not always the best. Especially, the markets average excess return rate is up to 6.488%, more than other portfolio, but risk is higher, and Sharpes ratio performance falls. Table 4 has all these detailed results.

In the third period (2008), market slowed down. Due to financial crisis, the performance of portfolio over the third period is not as good as that over the previous periods. In all, Sharpe value is positive. This shows portfolio in this study can still obtain return in declined market. Figure 6 reveals that the performance value of the four kinds of portfolio is better than the market return. For the number of winning months, HGP1 has 12 months.

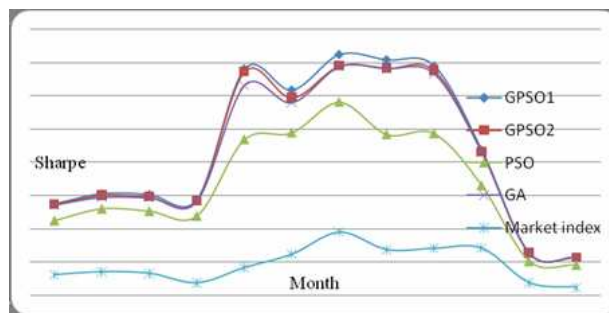


Fig. 5: Experimental results of second period (2007).

Table 4: Average monthly performance parameters of the second experiment (2007).

Portfolio	HGP1	HGP2	PSO	GA	Market Index
Average Sharpe value	2.200723	2.149749	1.679569	2.115209	0.475567
Excess return rate	3.839219	3.862478	4.389305	3.770062	6.488314
Total risk	1.744526	1.796711	2.613352	1.782359	13.6433333

For the average monthly Sharpe value over the third experimental period, HGP1 is the first, HGP2 is second, GA third, PSO fourth. For the number of winning months, GA is superior to PSO. See Table 5 for details.

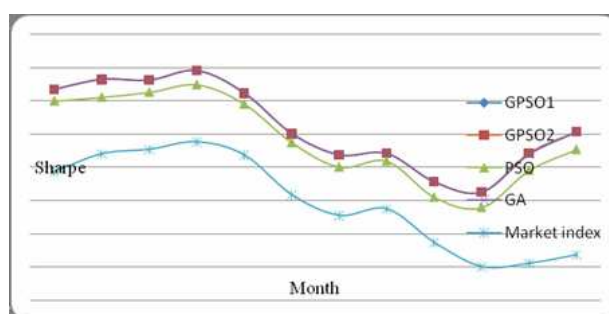


Fig. 6: Experimental results of third period (2008).

In the 30-time experiments per month, solution of each portfolio is different. Thus this study uses the standard deviation of the portfolios in the 36 months to disclose the stability. The results show that HGP1 has the best stability.

Table 5: Average monthly performance parameters of the second experiment (2008).

Portfolio	HGP1	HGP2	PSO	GA	Market Index
Average Sharpe value	0.251704	0.250647	0.170569	0.249638	-0.405533
Excess return rate	1.850726	1.848061	1.315119	1.840009	-10.99603
Total risk	7.352798	7.373161	7.710189	7.370707	27.115

4.6 Test of Hypothesis

It is noticed that Sharpe value based on HGP1 is greater than that in GA, PSO and HGP2 after evaluation of performance over 36 months. In order to validate whether Sharpe value bases on HGP1 is different from other algorithms, this study employs the Mann-Whitney U test with a confidence level of 95% to test whether average value of two independent samples based on HGP1 and HGP2, GA, PSO have difference. Take 01/2006 as an example (Table 6), the corresponding probability of HGP1 and other algorithms is 0.000, smaller than 0.05 of significant level. Thus, the null hypothesis H_0 is rejected, and it is considered that average values of HGP1 and other algorithms have significant difference. After integration of data over the three years, no significant months are depicted in Table 7.

Table 6: Validation of HGP1 and other algorithms (01/2006).

Algorithms	HGP1 vs. HGP2	HGP1 vs. GA	HGP1 vs. PSO
Mann-Whitney U	216.000	120.000	22.000
Wilcoxon W	681.000	585.000	487.000
Z	-3.494	-4.926	-6.390
Asymp. Sig. (2-tailed)	.000	.000	.000

Table 7: Validation of non-significant months in HGP1 and other three algorithms.

	2006	2007	2008
HGP2 portfolio	5, 6	11, 12	3, 5, 7, 11
HGP1 portfolio	0	0	0
GA portfolio	0	0	0

5 Conclusions

This study aims at fund selection based on DEA and asset allocation via hybrid of GA and PSO. It takes the Taiwan equity fund for validation. Performance of the four kinds of portfolio methods is all better than market return. According to validation analysis, HGP1 performance is significantly different from other algorithms, and HGP1 is proved to be the most effective. In the three experimental periods (2006, 2007, and 2008), portfolio performance based on HGP1 is the best and it is demonstrated that HGP1 has promising performance no matter what changes take place in market, like economic recession. The results also illustrate that average Sharpe value of four kinds of portfolio is better than market return. However, there is a gap in excess return rate, especially in 2007, market's excess return rate is much higher than the four kinds of portfolio. Hence, this study is not applicable for the investors seeking high return but for the investors who want to get gains in consideration of risk. The future study can introduce other soft computing techniques, like artificial immune system, to replace GA or PSO. Also, weight factor can be added to produce importance of variables.

References

- [1] Annaert J, Van Den Broeck J, Vennet R, Determinants of mutual fund underperformance: a Bayesian stochastic frontier approach. *European Journal of Operational Research*, **31**, 617-632 (2003).
- [2] Arumugam MS, Rao MVC, On the improved performances of the partial swarm optimization algorithms with adaptive parameters, cross-over operators and root mean square(RMS) vataiants for computing optimal control of a class of hybrid systems. *Applied Soft Computing*, **8**, 324-336 (2008).
- [3] Banker RD, Estimating most productive scale size using data envelopment analysis. *European Journal of Operational Research*, **17**, 35-44 (1984).
- [4] Banker RD, Charnes A, Cooper WW, Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, **30**, 1078-1092 (1984).
- [5] Basso A, Funari S, Measuring the performance of ethical mutual funds: a DEA approach selection. *Journal of the operational research society*, **54**, 521-531 (2003).
- [6] Basso A, Funari S, A data envelopment analysis approach to measure the mutual fund performance. *European Journal of Operational Research*, **135**, 477-492 (2001).
- [7] Basso A, Funari S, A generalized performance attribution technique for mutual funds. *Central European Journal of Operations Research*, **13**, 65-84 (2005).
- [8] Bauer JR, Richard J, Genetic algorithms and investment strategies. Wiley, 127-134 (1994).
- [9] Bergh FVD, EngelbrechtAP, A new locally convergent particle swarm optimizer. *Proceeding of IEEE International Conference on Systems, Man and Cybernetics*, 3-6 (2002).

- [10] Boeringer DW, Werner DH, Particle swarm optimization versus genetic algorithms for phased array synthesis. *IEEE Transactions on Antennas and Propagation*, 771-779 (2004).
- [11] Boyd R, Richerson PJ, Culture and the evolutionary process. University of Chicago Press, Chicago, (1985).
- [12] Brown Stephen J, Goetzman WN, Performance persistence. *The Journal of Finance*, 50, 1289-1307 (1995).
- [13] Chang JF, Hsu SW, The construction of stocks portfolios by using particle swarm optimization. *Proceedings of Innovative Computing, Information and Control*, 390-391 (2007).
- [14] Charnes A, Cooper WW, Rhodes E, Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429-444 (1978).
- [15] Chen SH, Yeh CH, Lee WC, Option pricing with genetic programming. *Third Annual International Genetic Programming Conference*, 22-25 (1988).
- [16] Chen W, Cai YM, Study on the efficient frontier in portfolio selection by using particle swarm optimization. *Chinese Control and Decision Conference*, 269-272 (2008).
- [17] Dashti MA, Farjami Y, Vedadi A, Anisseh M, Implementation of particle swarm optimization in construction of optimal risky portfolios. *Industrial Engineering and Engineering Management*, 812-816 (2007).
- [18] Dorigo M, Optimization, learning and natural algorithms. Ph.D. Thesis, Politecnico di Milano, (1992).
- [19] Droms WG, Walker DA, Performance persistence of international mutual funds. *Global Finance Journal*, 12, 237-248 (2001).
- [20] Du S, Li W, Cao K, A learning algorithm of artificial neural network based on GA-PSO. *Intelligent Control and Automation*, 1, 3633-3637 (2006).
- [21] Eberhart RC, Kennedy J, Particle swarm optimization. *Proceedings of the IEEE International Conference on Neural Network*, Perth, Australia, 1942-1948 (1995).
- [22] Eugene F, Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, 25, 383-417 (1970).
- [23] Fan SKS, Liang YC, Zahara E, A genetic algorithm and a particle swarm optimizer hybridized with Nelder-Mead simplex search. *Computer & Industrial Engineering*, 50, 401-425 (2006).
- [24] Grinblatt M, Titman S, The persistence of mutual fund performance. *Journal of Finance*, 47, (1992).
- [25] Hendry DF, C Santos, Regression models with data-based indicator variables. *Oxford Bulletin of Economics and Statistics*, 67, 571-595 (2005).
- [26] Holland J, Adaptation in natural and artificial system. University of Michigan Press, Ann Arbor, Michigan, USA, (1975).
- [27] Ippolito R A, Efficiency with costly informance: a study of mutual fund performance. *Quarterly Journal of Economics*, 104, 1-23 (1989).
- [28] Jensen M, The performance of mutual funds in the period 1945-1964. *Journal of Finance*, 23, 389-416 (1968).
- [29] Juang CF, A Hybrid of genetic algorithm and particle swarm optimization for recurrent network design. *IEEE Tran on System, Man, and Cybernetics part B: Cybernetics*, 34, 997-1006 (2004).
- [30] Kaboudan M, Using GP forecasts to enhance profitable trading of stocks. In *Proceedings of the 5th Joint Conference on Information Sciences*, 925-928 (2000).
- [31] Kao YT, Zahara E, A hybrid genetic algorithm and particle swarm optimization for multimodal function. *Applied Soft Computing*, 8, 849-857 (2007).
- [32] Khouja M, The use of data envelopment analysis for technology selection. *Computers and Industrial Engineering*, 28, 123-132 (1995).
- [33] Kim K, Han I, Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Systems with Applications*, 19, 125-132 (2006).
- [34] Kuo RJ, Chen JH, Hwang YC, An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network. *Fuzzy Sets and Systems*, 118, 21-45 (2001).
- [35] Laura D, Mihai O, Evolving the update strategy of the particle swarm optimization algorithms. *International Journal on Artificial Intelligence Tools*, 16, 87-109 (2007).
- [36] Lee JW, Kim SH, Using analytic network process and goal programming for interdependent information system project selection. *Computer and Operation Research*, 27, 367-382 (2000).
- [37] Maringer D, Kellerer H, Optimization of cardinality constrained portfolios with a hybrid local search algorithm. *OR Spectrum* 25, 481-495 (2003).
- [38] Markowitz HM, Portfolio selection. *Journal of Finance*, 7, 77-91 (1952).
- [39] McMullen, Patrick R, Robert A, Strong, Selection of mutual funds using data envelopment analysis. *Journal of Business and Economic Studies*, 4, 1-12 (1998).
- [40] Merwe DWVD, Engelbrecht AP, Data clustering using particle swarm optimization. *The Congress on Evolutionary Computation*, 215-220 (2003).
- [41] Murthi BPS, Yoon K, Choi, Preyas D, Efficiency of mutual funds and portfolio performance measurement: a non-parametric approach. *European Journal of Operational Research*, 98, 408-418 (1997).
- [42] Nikos ST, Timotheos A, Vassilios V, Georgios D, Active portfolio management with cardinality constraints: An application of particle swarm optimization. *January paper NMNC*, (2008).
- [43] Oh KJ, Kim TY, Min SH, Hyoung YL, Portfolio algorithm based on portfolio beta using genetic algorithm. *Expert Systems with Applications*, 30, 527-534 (2006).
- [44] Pareto V, *Manueldeconomic politique*, 2nd Edition, Girard, Paris, (1927).
- [45] Santos Andre, Joao T, Newton DC, Sergio DS, Evaluating brazilian eutual funds with stochastic frontiers abstract. *Economics Bulletin*, 13, 16 (2006).
- [46] Sarkis J, Evaluating flexible manufacturing systems alternatives using data envelopment analysis. *The Engineering Economist*, 43, 25-48 (1997).
- [47] Sharpe WF, Capital asset prices: a theory of market equilibrium under conditions of risk. *the Journal of Finance*, 19, 425-442 (1964).
- [48] Sharpe WF, Mutual fund performance. *Journal of Business*, 39, 119-138 (1966).
- [49] Shi Y, Eberhart R, A modified particle swarm optimizer. *Proceedings of the IEEE International Conference on Evolutionary Computatuin*, 69-73 (1998a).
- [50] Shi Y, Eberhart R, Parameter selection in particle swarm optimization. *Evolutionary Computatuin*, 591-600 (1998b).

- [51] Shi C, Lu J, Zhang G, An extended Kuhn-tucker approach for linear bi-level programming. *Applied Mathematics and Computation*, **162**, 51-63 (2005).
- [52] Smith KV, Tito DA, Risk-return measures of ex-post portfolio performance. *Journal of Financial and Quantitative Analysis*, **4**, 297-315 (1969).
- [53] Song G, Guo W, Chinese open-end fund operational efficiency appraisal using data envelopment analysis. *Proceedings of International Conference on Risk Management & Engineering Management*, 570-575 (2008).
- [54] Statman M, How many stocks make a diversified portfolio. *Journal of Financial and Quantitative Analysis*, **22**, 353- 363 (1987).
- [55] Tian J, Ma J, Study of security investment optimizing combination based on PSACO. *International Symposiums on Information Processing*, 710-714 (2008).
- [56] Treynor JL, How to rate management investment funds. *Harvard Business Review*, **43**, 63-75 (1961).
- [57] Wallin R, Ryan C, Maintaining diversity in EDAs for real-valued optimization problems. *Proceedings of the 2007 Frontiers in the Convergence of Bioscience and Information Technologies*, 795-800 (2007).
- [58] Wang GM, Wang XJ, Wang ZP, Chen YL, Genetic algorithm for solving learning bilevel programming. *Proceedings of the IEEE Sixth international Conference on Parallel and Distributed Computing*, 920-924 (2005).
- [59] Wermers R, Mutual fund performance: anempirical decomposition into stock-picking talent, style, transactions cost, and expenses. *Journal of Finance*, **55**, 1655-1703 (2000).
- [60] Xu F, Chen W, Yang, L, Improved particle swarm optimization for realistic portfolio selection. *Proceedings of the Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*, 185-190 (2007).
- [61] Zhao B, Guo CX, Bai BR, Cao YJ, An improved particle swarm optimization algorithm for unit commitment0. *International Journal of Electrical Power & Energy Systems*, **28**, 482-490 (2006).



R. J. Kuo received the MS degree in Industrial and Manufacturing Systems Engineering from Iowa State University, Ames, IA, in 1990 and the PhD degree in Industrial and Management Systems Engineering from the Pennsylvania State University, University Park, PA, in 1994. Currently, he is the Distinguished Professor in the Department of Industrial Management at National Taiwan University of Science and Technology, Taiwan. His research interests include architecture issues of computational intelligence and their applications to data mining, electronic business, logistics and supply chain management and decision support systems.



C. W. Hong received the MS degree in Industrial Engineering and Management from National Taipei University of Technology, Taiwan. Currently, he is the senior engineer for Innolux Corporation, Taiwan. His research interests include architecture issues of computational intelligence and their applications to data mining and operations management.