Proceedings of the World Congress on Engineering 2012 Vol I
WCE 2012, July 4 - 6, 2012, London, U.K.

Improved Portfolio Optimization Combining Multiobjective Evolutionary Computing Algorithm and Prediction Strategy

Sudhansu Kumar Mishra, Ganapati Panda, Babita Majhi, Ritanjali Majhi

Abstract— **In conventional mean-variance model of portfolio optimization problem the expected return is taken as the mean of the past returns. This assumption is not correct and hence the method leads to poor portfolio optimization performance. Hence an alternative but efficient method is proposed in which the mean and variance of expected return are first predicted with a low complexity functional link artificial neural network model (FLANN). The predicted values of mean and variance are consequently used in multi objective swarm intelligence techniques for achieving better performance. The multi objective swarm intelligence techniques chosen are non-dominated sorting genetic algorithm-II (NSGA – II) and multi objective particle swarm optimization (MOPSO).The performance of the proposed prediction based portfolio optimization model has been compared with the Markowitz mean-variance model. The comparison of the performance includes three performance metrics, Pareto front and nonparametric statistical test using the Sign test. On examining the performance metrics it is observed that the proposed prediction based portfolio optimization model approach provided improved Pareto solutions but maintaining adequate diversity.**

Index Terms— Constrained **Portfolio optimization, Efficient frontier, multiobjective optimization, Non-dominated sorting, nonparametric statistical test.**

I. INTRODUCTION

In the last few decades the portfolio optimization has emerged as a challenging and interesting multiobjective problem in the field of computational finance. It is receiving increasing attention of researchers, fund management companies and individual investor in the last few decades. Selecting a subset of assets and corresponding optimal weights from a set of available assets is a key issue in the portfolio optimization problem. The percentage of each available asset is selected in such a way that the total profit (return) of the portfolio is maximized while total risk to be minimized simultaneously. Harry Markowitz [1,2] set up a quantitative mean-variance framework for representing the risk and return of the portfolio, where the mean of past returns is taken as the expected return and the variance and covariance of the time series of return as the risk. Since the introduction of this mean-variance portfolio optimization model, considerable research attention has been made on model simplifications and the development of different risk measures. All these techniques use the meaning of the past return as expected return. Hence there is the need to develop efficient ways of approach which would directly predict the future return and the corresponding risk.

For the prediction of return the functional link artificial neural network (FLANN) has been used with its weight trained by evolutionary computing methods. The inputs to the network are some financial and economic variables which are selected by using evolutionary algorithms. The FLANN structure is used for prediction of return and the corresponding risk. Considering these two conflicting objectives the Portfolio optimization problem can be formulated as a multiobjective minimization problem and can be efficiently solved by using a multi objective evolutionary algorithm (MOEA). One of the main advantages of the MOEAs is that it gives a set of possible solutions in a single run which are known as a Pareto optimal solution [3, 4]. Pareto ant colony optimization (PACO) has been introduced for solving the portfolio selection problem in [4] and its performance has been compared with other heuristic approaches such as Pareto simulated annealing and the non-dominated sorting genetic algorithm.

In the present paper two multiobjective evolutionary algorithms based on non-dominating sorting such as NSGA-II and MOPSO have been suggested to solve the portfolio optimization problem using the proposed prediction based model. The performance of the proposed prediction based portfolio optimization model has been compared with the Markowitz mean-variance model using three performance metrics, Pareto front and nonparametric statistical test using the Sign test.

The rest of the paper is organized as follows. The multiobjective optimization is presented in a concise manner in Section 2. The two multiobjective evolutionary algorithms (MOEAs) such as NSGA-II and MOPSO which are used for portfolio optimization are discussed in Section 3. In Section 4 the proposed prediction based mean-variance model is described. Three performance metrics for assessing the performance of MOEAs are discussed in Section 5. Section 6

Manuscript received March 06, 2012,revised April 04, 2012
Sudhansu Kumar Mishra is with the Dept of ECE, National Institute of Technology, Rourkela, India(Sudhansu.nit@gmail.com)
Ganapati Panda is with the School of Electrical Sciences, Indian Institute of Technology, Bhubaneswar, India(ganapati.panda@gmail.com)
Babita Majhi is with the Dept. of CS & IT, G G Vishwavidayalaya, Central University Bilaspur, India (babita.majhi@gmail.com)
Ritanjali Majhi is with the School of Management, National Institute of Technology, Warangal, India(ritanjalimajhi@gmail.com)

provides the simulation results of present studies. Finally the conclusion of the investigation is presented in Section 7.

II. MULTIOBJECTIVE OPTIMIZATION: BASIC CONCEPTS AND BRIEF OVERVIEW

A multiobjective optimization problem (MOP) is defined as the problem of computing a vector of decision variables that satisfies the constraints and optimize a vector function whose elements represent the objective functions. The generalized multiobjective minimization problem may be formulated [4] as

Minimize $f\left(\vec{x}\right) = (f_1(\vec{x}), f_2(\vec{x}), \ldots\ldots\ldots, f_M(\vec{x}))$ (1)

Subjected to constraints:

$$g_j(x) \geq 0,\ j = 1,2,3\ldots..,J \quad (2)$$

$$h_k(x) = 0,\ k = 1,2,3\ldots\ldots..,K \quad (3)$$

where $\vec{x}$ represents a vector of decision variables $\vec{x} = \{x_1, x_{2,} \ldots\ldots., x_N\}^T$

The search space is limited by

$$x_i^L \leq x_i \leq x_i^U,\ i = 1,2,3\ldots\ldots.,N \quad (4)$$

x_i^L and x_i^U represent the lower and upper acceptable values respectively for the variable x_i.

N represents the number of decision variables and M represents the number of objective functions. Any solution vector $\vec{u} = \{u_1, u_2, \ldots\ldots u_K\}^T$ is said to dominate over $\vec{v} = \{v_1, v_2, \ldots\ldots, v_k\}^T$ if and only if

$$f_i\left(\vec{u}\right) \leq f_i\left(\vec{v}\right)\ \forall i \in \{1,2,\ldots..,M\}$$

$$f_i\left(\vec{u}\right) < f_i\left(\vec{v}\right)\ \exists i \in \{1,2,\ldots\ldots.,M\} \quad (5)$$

Those solutions which are not dominated by other solutions for a given set are considered as non-dominated solutions. The front obtained by mapping these non-dominated solutions into objective space is called Pareto-optimal front (POF).

The generalized concept of Pareto front was introduced by Pareto in 1986 [5]. The practical application of genetic algorithm to multiobjective optimization problem (MOP) such as vector evaluated genetic algorithm (VEGA) [6], SPEA-2 [7], NSGA-II [8] etc. have been proposed by many authors. In recent past heuristic approach based on particle swarm optimization to solve multiobjective problems has been introduced by Coello et al. [9]. Some approach based on particle swarm optimization such as TV-MOPSO [10], FCPSO [11] etc. have been suggested to solve the MOP. In this paper two competitive MOEAs (NSGA-II and MOPSO) which have already been applied efficiently to the portfolio optimization problem using Markowitz mean variance model [12] is applied using proposed mean variance model.

III. MULTIOBJECTIVE SWARM INTELLIGENCE TECHNIQUES FOR PORTFOLIO OPTIMIZATION

The classical optimization techniques are ineffective for solving constrained optimization problem such as portfolio management [3]. This shortcoming has motivated the researchers to develop multi-objective optimization using evolutionary techniques. In this paper we have compared the portfolio optimization performance achieved by of two recently developed multi-objective evolutionary algorithms such as NSGA-II [8] and MOPSO [9] by using our proposed and Markowitz mean-variance model. When these MOEAs are applied for portfolio optimization, issues like representation, variation operator and constraint handling techniques are considered. The NSGA-II maintains a population of chromosome, where each of them represents a potential solution to the portfolio optimization problem. One chromosome represented by a weight vector, provides the composition of the portfolio. In MOPSO the position of each particle represents a weight vector associated (percentage) with different assets.

IV. THE PROPOSED PREDICTION BASED MEAN-VARIANCE MODEL

This section proposes a prediction-based portfolio optimization model. This model uses predicted returns as expected returns instead of using the mean of past returns. Also instead of using the variance of the returns it uses the variance of the errors of the predicted return as risk measure. An investment is planned over a time period and its performance is measured using its return that quantifies the wealth variation. The one period stock return in time t is defined as the difference between the price of the stock at time t and the price at time $t-1$, divided by the price at time $t-1$. Mathematically it is expressed as:

$$R_t = (P_t - P_{t-1}) / P_{t-1}\ ,\ t \geq 1 \quad (6)$$

where R_t is the one-period stock return at time t, and P_t and P_{t-1} are the stock prices at times t and $t-1$, respectively. The series of N past returns of a stock, R', is defined as

$$R = (R_1, R_2, \ldots.., R_N) \quad (7)$$

The one-period prediction of the future return of a stock can be defined as the process of using R' for obtaining an estimate of R_{t+1}.

Different variants of neural network have been applied for prediction of stock indices [13]. In this paper an efficient single layer neural network called as functional link artificial neural network (FLANN) is used for prediction which is trained with evolutionary computing. The inputs to the network are some financial and economic variables which are judiciously selected by using evolutionary algorithms. Pao originally proposed FLANN and it is a novel single layer ANN structure capable of forming arbitrarily complex decision regions by generating nonlinear decision boundaries [14]. Here, the input is enhanced by using nonlinear function. This nonlinear functional expansion of the input pattern may be trigonometric, exponential, power series or Chebyshev type. The architecture of FLANN is simple and its training can be performed using standard steepest descent or evolutionary computing algorithm. The prediction of stock return is a nonlinear task and can be conventionally performed using a FLANN structure.

Let the predicted and actual return be related may be represented as

$$R_t = \hat{R}_t + E_t \tag{8}$$

where R_t and $\hat{R}_t$ be the actual stock return and predicted stock return at time t respectively and E_t is the prediction error at time t and is defined as

$$E_t = R_t - \hat{R}_t \tag{9}$$

The time series of n errors of prediction may be represented as

$$E = (E_1, E_2, \ldots\ldots, E_n) \tag{10}$$

A portfolio is a collection of N stocks and N weights, or participations. The participation, of each asset is w_i, $i = 0,1,2.....N.$ Where $0 \le w_i \le 1$ represents the fraction of the portfolio value invested in the stock i such that

$$\sum_{i=1}^{N} w_i = 1 \tag{11}$$

It shows the budget constraint which ensures that the sum of the weights associated with each asset is equal to one which means all the available money is invested in the portfolio. The predicted return of the portfolio, or portfolio expected return, R_p, is the participations and predicted returns of the stocks of the portfolio.

$$R_p = \sum_{i=1}^{N} w_i \hat{R}_i \tag{12}$$

The portfolio risk is the variance of the joint normal distribution of the linear combination of the participations and prediction errors of the stocks of the portfolio

$$V = \hat{\sigma}_p^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \gamma_{Eij} \tag{13}$$

where $\hat{\sigma}_p^2$ is the total portfolio risk and is equal to the variance of the linear combination of the participations and prediction errors of the stocks of the portfolio. N is the number of stocks in the portfolio. w_i and w_j are the participations of stocks i and j of the portfolio respectively. γ_{Eij} is the interactive prediction risk of stocks i and j, which is the covariance of the errors of prediction of the stocks i and j.

The prediction-based portfolio optimization model can also be formulated as single objective optimization as

Minimize $\lambda\left[\hat{\sigma}_p^2\right] - (1-\lambda)\left[R_p\right]$

$$= \lambda\left[\sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \gamma_{Eij}\right] - (1-\lambda)\left[\sum_{i=1}^{N} w_i \hat{R}_i\right] \tag{14}$$

Hence such a formulation yields non-dominated solutions by varying the $\lambda(0 \le \lambda \le 1)$ factor. But in the present case the multiobjective portfolio optimization problem is solved by using two MOEAs. It does not combine the two objectives to obtain the Pareto optimal solution set. Here the two objectives are taken individually and the algorithm tends to optimize both of them simultaneously. In the proposed work the two objectives are expressed as minimization objective. To express both the objectives in minimization form, the second objective R_p is expressed as $-R_p$. Accordingly the portfolio problem is expressed as minimization of $\hat{\sigma}_p^2$ and $-R_p$ simultaneously.

VI. SIMULATION STUDY

The algorithms are coded in MATLAB and are run on a PC with Intel Core2 Duo 3.0 GHz with 4 GB RAM. The portfolio optimization problem is solved using the prediction based mean-variance model by applying NSGA-II and NS-MOPSO algorithms. The results are compared with the Markowitz mean-variance model by applying the same two MOEAs algorithms. For the two MOEAs the population size and number of generations are taken as 50 and 1000 respectively. In MOPSO, the position of each particle represents a weight vector associated (percentage) with different assets. Where as in NSGA-II, one chromosome represents one set of weights of assets and each gene represents weight of one asset. After conducting several experiments with different parameters, the final parameters for the two algorithms are chosen as

NSGA-II: The uniform crossover and mutation rates are taken as 0.08 and 0.05 respectively.

MOPSO: The velocity probability is taken 0.5 in a different direction. Its upper and lower bounds, V_{UPP} and V_{LOW}, are fixed at 0.06 and 0.5 respectively. The parameter $w = 0.862$ and $C_1 = C_2 = 2.05$.

ISBN: 978-988-19251-3-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Two different metrics defined in the sequel are used during the investigation for measuring the performance quality are:
Generation distance (GD): It estimates the distance of elements of non-dominated vectors found, from those standard efficient frontiers and is mathematically [15] expressed as

$$GD = \frac{\sqrt{\sum_{i=1}^{n} d_i^{\,2}}}{n} \tag{15}$$

where n is the number of vectors in the set of non-dominated solutions. d_i is the Euclidean distance between each of these and the nearest member of the standard efficient frontier. If $GD = 0$, all the candidate solutions are in standard efficient frontier. The smaller the value of GD the closer is the solution to the standard efficient frontier.
Spacing (S): It measures the spread of candidate solution throughout the non-dominated vectors found. This metric [15] is mathematically expressed as

$$S \stackrel{\Delta}{=} \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\bar{d} - d_i\right)^2} \tag{16}$$

where

$$d_i = \min_j \left(\left| f_1^{\,i}\left(\vec{x}\right) - f_1^{\,j}\left(\vec{x}\right) \right| + \left| f_2^{\,i}\left(\vec{x}\right) - f_2^{\,j}\left(\vec{x}\right) \right| \right)$$

and $i, j = 1,2,...,n$

$\bar{d} =$ mean of all d_i and n is the number of non-dominated vectors found so far. A value of zero for this metric indicates all members of the Pareto front currently available are spaced at equidistant.

All the experiments have been conducted with a set of benchmark data available online and obtained from OR-Library [16]. The data corresponds to weekly prices between March 1992 and September 1997 from different well known indices such as Hang-Seng in Hong Kong, DAX 100 in Germany, FTSE100 in UK, S&P 100 in USA and Nikkei225 in Japan. The numbers of different assets for the above benchmark indices are 31, 85, 89, 98 and 225 respectively. Using each data the mean return of individual assets is calculated from the weekly price. In the data set the correlation between assets are also given. The covariance between the assets, evaluated from the correlation matrix, is used for calculating the risk of portfolio. But in the proposed mean-variance model we have not used the calculated mean return rather applied our proposed FLANN based predictor to predict the return. The corresponding risk is calculated and finally utilized by MOEAs algorithms for risk-return tradeoff.

The data (risk and corresponding tradeoff return) for standard efficient frontiers for the five stocks are represented by PORTEF-1 to PORTEF-5 and are found at http://people.brunel.ac.uk/~mastjjb/jeb/orlib/portinfo.html [16]. PORTEF-1 to PORTEF-5 belong to Hang -Seng, DAX 100, FTSE 100, S&P 100 and Nikkei 225 stock indices respectively. In our study we have used the Nikkei 225 stock having more number of assets to test our proposed technique. The Pareto front corresponding to Hang-Seng stock indices i.e. PORTEF-1, called as standard efficient front or global optimal Pareto front (GOPF) is depicted in Fig.1. It shows the tradeoff between risk (variance of return) and return (mean return).

The two MOEAs are also applied to Hang-Seng stock indices using both the proposed and Markowitz mean-variance model. The corresponding Pareto front obtained and the standard efficient frontiers are shown in Fig.2. It is evident that the two algorithms are capable of providing good solutions using the proposed mean-variance model.

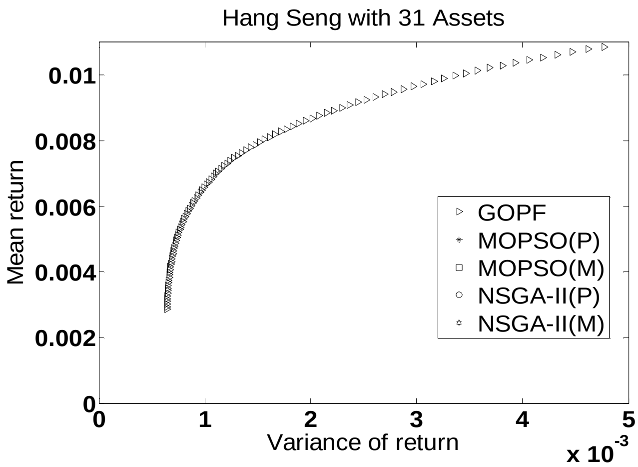

Fig.1 Standard efficient front for Hang-Seng stock Indices

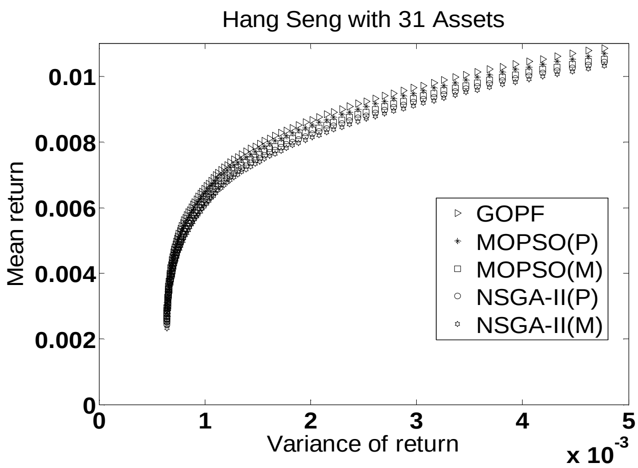

Fig.2 The Standard efficient front and Pareto front of other algorithms

Further, the performance of the two MOEAs is assessed by applying it to Hang-Seng stock and using two different metrics such as the GD and S. The two algorithms are run for 25 times and then the maximum, minimum, average and standard deviation of the two metrics are calculated and the corresponding results are shown in Table-1.
M: Markowitz model, P: Proposed model

ISBN: 978-988-19251-3-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE I
Comparison of performance evaluation metrics obtained using different MOEAs

Algorithm		NSGA-II (M)	NSGA-II (P)	MOPSO (M)	MOPSO (P)
S	Max	9.21E-3	7.43E-3	6.54E-3	5.12E-3
	Min	7.87E-3	5.23E-3	4.98E-6	3.88E-6
	Avg	8.33E-3	6.36E-3	5.74E-3	5.13E-3
	Std.	2.58E-3	1.58E-3	1.53E-3	1.23E-3
GD	Max	2.54E-2	2.01E-2	1.98E-2	1.34E-2
	Min	1.01E-2	0.89E-2	0.78E-2	0.56E-2
	Avg	1.76E-2	1.02E-2	0.87E-2	0.79E-2
	Std.	0.42E-2	0.28E-2	0.22E-2	0.21E-2

The results demonstrate that the proposed model is better for stock having a small number of assets such as Hang-Seng, DAX 100, FTSE 100, S&P 100 benchmark indices having 31, 85, 89, 98 different assets. The convergence (C) metrics for two MOEAs for Hang-Seng stock are demonstrated in Table-2. It is found that the most of the solutions obtained by MOPSO algorithm with predicted mean-variance model dominate the solutions obtained from other three cases. The nonparametric statistical test such as Sign test is carried out for pairwise comparisons of the performance of two algorithms [17]. The results of the Sign test for pairwise comparisons among proposed NS-MOPSO and other algorithms while taking the S metric as the wining parameter (i.e. lower value of S means win) are shown in Table-3. From the results it is concluded proposed MOPSO and predicted mean variance model shows improved performance compared to its counterpart.

TABLE II
Comparison of results of C metric obtained using different MOEAs

	NSGA-II (M)	NSGA-II (P)	MOPSO (M)	MOPSO (P)
NSGA-II(M)	—	0.3810	0.2230	0.1781
NSGA-II(P)	0.4120	—	0.2580	0.2168
MOPSO(M)	0.8520	0.7630	—	0.3244
MOPSO(P)	0.9078	0.8210	0.3354	—

TABLE III
Critical values for the two-tailed Sign Test at $\alpha = 0.05$ and $\alpha = 0.1$ .using S metric as winning parameter.

MOPSO(P)	NSGA-II(M)	NSGA-II(P)	MOPSO(M)
Wins(+)	20	17	13
Loses(-)	5	8	12
Detected differences	$\alpha = 0.05$	$\alpha = 0.01$	--

VII. CONCLUSION

A novel prediction based portfolio optimization model has been proposed and two multiobjective evolutionary algorithms NSGA-II and MOPSO have been employed to solve the portfolio optimization problem. In the proposed method the return is predicted with a low complexity single layer neural network. The performance of the proposed prediction based portfolio optimization model and the Markowitz mean-variance model has been evaluated and compared using two performance metrics. In addition to this, in the present study a Sign test [17] is carried out to pairwise compare the performance of the algorithms. From the simulation results it is observed that the proposed prediction based portfolio optimization model is capable of identifying good Pareto solutions maintaining adequate diversity and the performance is comparable with the well known Markowitz mean variance model. Further study in this field may include performance evaluation of the MOEAs using the proposed model considering some real world constraints like cardinality, ceiling, floor, round-lot, turnover etc. The same multiobjective optimization algorithm can also be applied to other financial applications such as asset allocation, risk management and option pricing.

REFERENCES

[1] H .M. Markowitz. Portfolio Selection, Journal of Finance. 1952; 7: 77-91
[2] H.M. Markowitz. Portfolio Selection: efficient diversification of investments. New York: Yale University Press. John Wiley & Sons, 1991.
[3] A. Mukherjee, R. Biswas, K. Deb and A. P. Mathur. Multi-objective Evolutionary algorithms for the Risk-return trade-off in bank loan management, International transaction in operations research, 2002; 9: 583-597.
[4] K. Doerner, W. Gutjahr, R. Hartl, C. Strauss, C. Stummer. Pareto ant colony optimization: A metaheuristic approach to multiobjective portfolio selection. Annals of Operations Research, 2004; 131: 79–99.
[5] Vilfredo Pareto, D. Cours. EconomiePolitique. Volume.8. F. Rouge, Lausanne 1986.
[6] J.D.Schaffer. Multiple objective optimizations with vector evaluated genetic algorithms. In Genetic Algorithms and their Applications: Proceedings of the International Conference on Genetic Algorithm, Lawrence Erlbaum 1985; 93-100.
[7] E. Zitzler, M. Laumanns, L. Thiele. SPEA2: Improving the strength Pareto evolutionary algorithm. Technical Report 103, Gloriastrasse 35, CH-8092 Zurich, Switzerland, 2001.
[8] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computing 2002; 6 (2): 182–197.
[9] C.A.C.Coello, T.P.Gregorio. Maximino.S.L. Handling Multiple objectives with particle swarm optimization, IEEE transaction on evolutionary computation, 2004; 8(3): 256-279.
[10] P.K. Tripathy, S. Bandyopadhyay, S. K. Pal. Multi-objective particle swarm optimization with time variant inertia and acceleration coefficients. Information Sciences, 2007; 177(22): 5033-5049.
[11] Shubham Agrawal. B.K. Panigrahi, M.K. Tiwari. Multiobjective particle swarm algorithm with fuzzy clustering for electrical power dispatch, IEEE Trans Evolutionary Computing 2008; 12 (5): 529–541.
[12] S.K. Mishra, G. Panda and S. Meher, Multiobjective particle swarm optimization approach to portfolio optimization, IEEE World Congress on Nature and Biologically Inspired Computing (NaBIC), Coimbatore, India 2009;1612-1615.
[13] J.S.Ziurilli, Financial prediction using neural networks, International Thompson Computer Press, London, 1997.
[14] Y.H. Pao, Adaptive Pattern Recognition and Neural Networks, Addison Wesley, Reading, Massachusetts, 1989.
[15] E. Zitzler, Evolutionary algorithms for multiobjective optimization: Methods and applications, Doctoral dissertation ETH 13398, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, 1999.
[16] http://people.brunel.ac.uk/~mastjjb/jeb/orlib/portinfo.html
[17] J. Derrac, S. García, D. Molina, F. Herrera, "A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms", Swarm and Evolutionary Computation,2011; 1(1): 3-18.

ISBN: 978-988-19251-3-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)