International Journal of Recent Trends in Engineering, Vol 2, No. 5, November 2009

Optimal Weighting of Assets using a Multi-objective Evolutionary Algorithm

Sudhansu Kumar Mishra1, Ganapati Panda2, Sukadev Meher3, sitanshu sekhar sahu4
Department of Electronics & Communication, NIT Rourkela, India-769008,
sudhansu.nit@gmail.com, ganapati.panda@gmail.com, smeher@nitrkl.ac.in,sitanshusekhar@gmail.com

***Abstract*— The problem of portfolio optimization is a well-known standard problem in financial world. It has received a lot of attention among many researchers. Choosing an optimal weighting of assets is a critical issue for which the decision maker takes several aspects into consideration. In this paper we consider a multi-objective portfolio assets selection problem where the total profit of is maximized while total risk to be minimized simultaneously. Three well-known multiobjective evolutionary algorithms i.e. Pareto Envelope-based Selection Algorithm(PESA), Strength Pareto Evolutionary Algorithm 2(SPEA2), Nondominated Sorting Genetic Algorithm II(NSGA II) for solving the bi-objective portfolio optimization problem has been applied. Performance comparison carried out in this paper by performing different numerical experiments. These experiments are performed using real-world data. The results show that NSGA-II outperforms other two for the considered test cases.**

***Index Terms*—Genetic algorithms, multiobjective optimization, Pareto-optimal solutions. global optimization, Crowding distance.**

INTRODUCTION

Choosing an optimal portfolio weighting of assets is main aim of portfolio optimization problem. Portfolio optimization is very complicated as it depends on factors like assets interrelationships, preferences of the decision makers, resource allocation and several other factors. Selecting an optimal portfolio weighting of assets, when their future rate of return is uncertain can be seen as a problem of minimizing the uncertainty for a given level of the portfolio expected return. The risk of a particular investment is not as important as its contribution to total portfolio risk. Combining a riskful investment with one carrying less risk it is possible to reduce the total risk associated to that portfolio.

Markovitz proposed the "expected return - variance (E-V)" model: maximizing the expected return for a unit of assumed risk [4]. Markovitz modern portfolio theory has originated from the idea that the investor is interested in two fundamental aspects i.e. risk and return. The risk of a particular investment is not as important as its contribution to total portfolio risk. Combining a riskful investment with one carrying less risk it is possible to reduce the total risk associated to that portfolio. Therefore, In this paper we suggest the use of multiobjective optimization algorithms for optimal weighting of assets as a portfolio optimization problem. In this paper we use PESA, SPEA2 and NSGA II for modeling the Pareto front and for optimizing the portfolio performance. The results obtained with these algorithms are finally compared by performing different numerical experiments. Selecting an optimal portfolio weighting of assets, when their future rate of return is uncertain is seen as a problem of minimizing the uncertainty for a given level of the portfolio expected return. This uncertainty is called as risk and measured by standard deviation of the probability distribution of future return.

Portfolio p consisting of N assets. Selection of optimal weighting of assets (with specific volumes for each asset given by weights w_i) is to be found. The unconstrained portfolio optimization problem is given as minimizing the variance of the portfolio and maximizing the return of the portfolio as defined in equation 1 and 2 respectively.

$$\rho_p{}^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{ij} \qquad (1)$$

$$\alpha_p = \sum_{i}^{N} w_i \mu_i \qquad (2)$$

$$\sum_{i}^{N} w_i = 1 \qquad (3)$$

$$0 \le w_i \le 1; \text{ and } i = 1,2\ldots,N \qquad (4)$$

Where N is the number of assets available, μ_i the expected return of asset i, σ_{ij} the covariance between asset i and j, and finally w_i are the decision variables giving the composition of the portfolio. ρ_p be the standard deviation of portfolio

International Journal of Recent Trends in Engineering, Vol 2, No. 5, November 2009

and α_p be the expected return of portfolio. This is a multiobjective optimization problem with two competing objectives. First is to minimize the variance (risk) of the portfolio and at the same the return of the portfolio will be maximized. The last two equation 3 and 4 gives the constraints for this portfolio optimization problem. Section II briefly describes the multi-objective optimization problem. In Section III some of the multi-objective evolutionary techniques are shortly described. Simulation studies based on several numerical experiments based on related work are performed in section IV. The Convergence characteristics are shown in section V. Conclusions and further research work directions are discussed in the section VI.

II. MULTI-OBJECTIVE OPTIMIZATION

In a single-objective optimization problem, an optimal solution is the one which optimizes the objective with certain model constraints. It is not possible to find a single solution for a multiobjective problem and due to the contradictory objectives a set of solutions are found. The general multi-objective minimization problem is formulated as: In multi-objective optimization, we are tasked with minimizing not just one objective function, but n objective functions:

$\left\{ f_1\left(\bar{x}\right), f_2\left(\bar{x}\right)........., f_n\left(\bar{x}\right)\right\}$ where $n \geq 2$

The solution to this problem is more complex than the single-objective case, and the idea of Pareto-dominance must be introduced to be able to visualize it. Consider first an objective function $\bar{F}\left(\bar{x}\right)$, where

$$\bar{F}\left(\bar{x}\right) = \left\{ f_1\left(\bar{x}\right), f_2\left(\bar{x}\right)........., f_n\left(\bar{x}\right)\right\} \quad (5)$$

A point $\bar{x}_1$ with an objective function vector $\bar{F}_1$, is said to dominate point $\bar{x}_2$, with an objective function vector $\bar{F}_2$, if no component of $\bar{F}_1$ is greater than its corresponding component in $\bar{F}_2$, and at least one component is smaller. Similarly, $\bar{x}_1$ can be said to be Pareto-equivalent to $\bar{x}_2$ if some components of $\bar{F}_1$ are greater than $\bar{F}_2$ and some are smaller. Pareto-equivalent points represent a trade-off between the objective functions, and it is impossible to say that one point is better than another Paretoequivalent point without introducing preferences or relative weighting of the objectives.

Thus, the solution to a multi-objective optimization problem is a set of design vectors which are not dominated by any other vector, and which are Pareto- equivalent to each other. This set is known as the Pareto-optimal set.

Some basic multi-objective concepts are:

1.Non-dominated:All decision vectors which dominate others but do not dominate themselves are called non-dominated or optimal solutions in the Pareto sense.

2.Local optimality: A vector u is locally optimal in the Pareto sense, if there exists a real $e > 0$ such that there is no vector $u\left(\bar{x}\right)$ which dominates the vector u with $u\left(\bar{x}\right)$.

3.Global optimality: A vector u is globally optimal in the Pareto sense, if there does not exist any vector $u\left(\bar{x}\right)$ such that vector $u\left(\bar{x}\right)$ dominates the vector u.

4.Pareto-optimal set : The set of all globally optimal solutions is called the Pareto-optimal set, and the set of all non-dominated objective vectors is called the Pareto Front (PF).

5.Pareto optimization: Finding an approximation to either the Pareto-optimal set or the Pareto front is referred to as Pareto optimization.

Two objectives are considered: maximizing the total profit and minimizing the total risk of the portfolio by selecting different weighting of total available assets. These two objectives are contradictory i.e to increase profit, one have to take more risk. Interdependencies are considered in the profit objectives while calculating risk interdependencies is rather difficult and not very meaningful since risk is derived as an average of several dimensions. This paper describes a situation in which 31 assets are available and an optimal portfolio weighting of these assets are needed. We assume that interdependencies exist among these assets.

III. MULTIOBJECTIVE EVOLUTIONARY ALGORITHMS

Multi-objective evolutionary algorithms are popular approaches in dealing with problems which consider several objectives to optimize. In this paper we compared the performance of three recently developed multiobjective evolutionary algorithms that are: Pareto Envelope-based Selection Algorithm, Strength Pareto Evolutionary Algorithm 2, Nondominated Sorting Genetic Algorithm II for optimal weighting of assets in portfolio optimization problem.

In PESA mating selection procedure is based on a crowding measure. The crowding distance measurement is done over the archive members. Crowding strategy works by forming hyper-grid and it divides phenotype space into hyper-boxes. Each individual in the archive is associated with a particular hyper-box. It has a squeeze factor which is equal to the number of other individuals from archive which present in the same hyper box. Environmental selection criteria based on this crowding measure is used for each individuals from archive . The main attraction of PESA is the integration of selection and diversity maintenance, whereby essentially the same technique is used for both tasks.

In SPEA 2 mating selection procedure is based on fitness measure and it uses binary tournament operator. Here the archive update is performed according to the fitness values associated with each of the individual in the archive. All individuals that have fitness less than 1 fill the archive and if the archive size is less than pre-established size, the archive is completed with dominated individuals from current pool. If the archive size exceeds the pre-established size, some individuals are removed from archive using the truncation operator. This operator is based on the distance of an individual to its nearest neighbor.

In NSGA II mating selection is based on the ranking criteria just like PESA and selection criteria is based on the crowding comparison operator.. Here the pool of individuals is split into different fronts and each front has assigned a specific rank. All individuals from a front F_i are ordered according to a crowding measure which is equal to the sum of distance to the two closest individuals along each objective. The environmental selection is processed based on these ranks. The archive will be formed by the nondominated individuals from each front and it begins with the best ranking front. Here the new population obtained after environmental selection is used for selection crossover and mutation to create a new population. It uses a binary tournament selection operator.

A. PESA Algorithm.

PESA has two parameters concerning population size i.e. p_I (the size of internal population, IP) and PE (the maximum size of external population EP) . It has one parameter concerning the hyper-grid crowding strategy.

1. Generate and evaluate each of an initial internal population (IP) of PI chromosomes.
2. Initialize the external population (EP) as empty set.
3. For $t = 1$ to Number of Generations
3.1. Incorporate the non-dominated members of IP into EP.
3.2. Delete the current content of IP.
3.3. Until obtain new solution of p_I .
3.3.1.Select two parents from EP with probability p_c
3.3.2. Recombination this two parents for obtaining one offspring
3.3.3. Mutate the offspring
3.3.4. Select one parent from IP with probability $(1 - p_c)$
3.3.5. Mutate the parent to produce one offspring
3.3.6. Add the two obtained offspring into IP
4. Return to 3

B. The SPEA2 Algorithm

Input: N (population size)
$\bar{N}$ (archive size)
T (maximum number of generations)
Output: A (nondominated set)
SPEA2 algorithm has five main steps i.e initialization, fitness assignment, environmental selection, termination and mating selection.
1. Generate and Initialize Population
2. Create empty external set $\bar{P}_0$ i.e the archive
3. Evaluate fitness values of each individual in P_0
4. For t = 1 to Number of Generations
4.1 Calculate fitness of each individual in P_t and $\bar{p}_t$
4.2 Copy all nondominated individuals in P_t and $\bar{p}_t$ to $\bar{P}_{t+1}$
4.3 If the $\bar{P}_{t+1}$ size exceeds archive size (pre-established) reduce $\bar{P}_{t+1}$ using truncation operator

4.3 If the $\bar{P}_{t+1}$ size is less than archive size then use dominated individuals in P_t to fill $\bar{P}_{t+1}$

4.4 Perform Binary Tournament Selection with replacement on $\bar{P}_{t+1}$ to fill the mating pool

4.5 Apply crossover and mutation to the mating pool and update P_t

4.6 Return to 4

C. NSGA II Algorithm

1. Initialize Population
2. Generate random Parent Population p_0 of size N
3. Evaluate Objective Values
4. Assign Fitness (or Rank) equal to its nondominated level
5. Generate Offspring Population Q_0 of size N with Binary Tournament Selection, Recombination and Mutation
6. For $t = 1$ to Number of Generations

6.1 Combine Parent and Offspring Populations

6.2 Assign Rank (level) based on Pareto Dominance.

6.3 Generate sets of non-dominated fronts

6.4 until the parent population is filled do

6.4.1 determine Crowding distance between points on each front F_i

6.4.2 include the ith nondominated front in the next parent population (P_{t+1})

6.4.3 check the next front for inclusion

6.5 Sort the front in descending order using Crowded comparison operator

6.6 Choose the first N - card (P_{t+1}) elements from front and include them in the next parent population (P_{t+1})

6.7 Using Binary Tournament Selection, Recombination and Mutation Create next generation

7. Return to 6

IV. Simulation Studies

In this section we describe the test problem used to compare the performance of PESA, SPEA 2, and NSGA II and for optimal weighting of the available assets. In all the cases the objective number is 2. We took parameters of these algorithms such a way that it will be comparable. We run experiments on data from OR library that maintained by Prof. Beasley as a public benchmark data set and is derived from Heng Seng data set with 31 assets. The data can be found at http://people.brunel.ac.uk/~mastjjb/jeb/orlib/portinfo.html.

The PESA has internal population size of 100, external population size of 100 and number of gene is equal to number of assets. We took number of generations as 100, uniform crossover rate as 0.8 and mutation rate 0.05. The grid size i.e. the number of division per dimension is 10. The NSGA II has population size of 100, number of generations 100, Crossover rate 0.8 and mutation rate 0.05. The numbers of real-coded variables are equal to number of assets and the selection strategy was tournament selection. The SPEA2 has population size of 100, number of generation be 100, crossover rate 0.8 and mutation rate 0.05. The gene length is equal to number of assets.

1. S metric

The S metric how much of the objective space is dominated by a given nondominated set A. If the S metric of a nondominated front f_1 is less than another front f_2 then f_1 better then f_2. It has been proposed by Zitzler[8].

2. Δ *metric*

This metric called as spacing metric (Δ) measures how evenly the points in the approximation set are distributed in the objective space. This formulation for this metric is given by:

$$\Delta = \frac{d_f + d_l + \sum_{i=1}^{N-1} \left| d_i - \bar{d} \right|}{d_f + d_l + (N-1)\bar{d}} \qquad (6)$$

Where d_i be the Euclidean distance between consecutive solutions in the obtained nondominated set of solutions. $\bar{d}$ is the average of these distances. d_f and d_l are the Euclidean distance between the extreme solutions and the boundary solutions of the obtained nondominated set and N is the number of solutions from nondominated set. The low value for Δ indicate a better diversity and hence better the algorithm.

3. C metric

Two sets of nondominated solutions can be compared Using C metric. For any two set A and B the C metric will be:

$$C(A,B) = \frac{|\{b \in B \mid \exists a \in A : a \succ b\}|}{|B|} \quad (7)$$

TABLE I
THE S AND Δ METRICS

Algorithm	PESA	SPEA2	NSGA II
Metric S	0.000304616	0.0000067874	0.000000574
Metric Δ	0.865412859	0.8337976192	0.5967844252

Table I shows the S metric and Δ metric obtained using all the three algorithms. NSGA II performs better as both S and Δ metric values are less than other two algorithms.

TABLE VI
THE RESULTS OBTAINED FOR C METRIC

	PESA	NSGA II	SPEA2
PESA	—	0.0000	0.0000
NSGA II	0.95790	—	0.2566
SPEA2	0.94627	.08534	—

The values 0.9579 on the second line, first column means almost all solutions from final populations obtained by NSGA II dominate the solutions obtained by PESA. The values 0 on first row means that no solution from the nondominated population obtained by SPEA 2 and by NSGA II is dominated by solutions from final populations obtained by PESA.
The performance of the two algorithms NSGA II and SPEA 2 are almost same but closely analyzing the value of C we can conclude that NSGA II performance better than SPEA 2.

V. CONVERGENCE CHARACTERISTICS

The Pareto front generated by these three algorithms is depicted as:

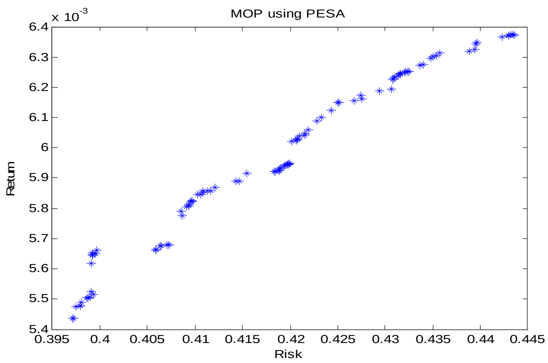

Figure 1. PESA

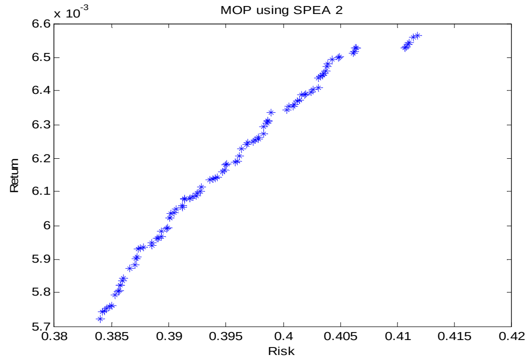

Figure2. SPEA 2

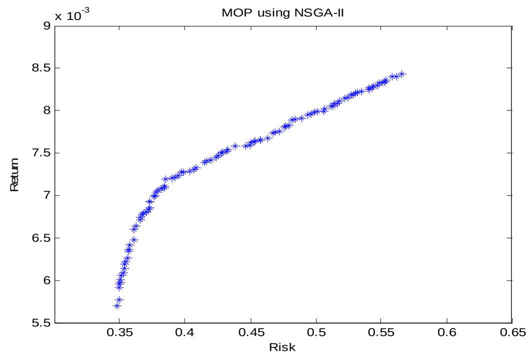

Figure 3. NSGA II

VI. CONCLUSIONS AND FURTHER WORK

In this paper a comparison of 3 multi-objective evolutionary algorithms for solving efficient weighting of assets portfolio optimization problem has been performed. The compared algorithms are PESA, SPEA2 and NSGA2. An assets set is considered for the numerical experiments. Results have shown that the NSGA II significantly outperforms the compared algorithms in all experiments.

REFERENCES

[1]K. Deb, A. Pratab, S. Agrawal, and T. Meyarivan. "A Fast Elitist Non-Dominated Sorting Genetic Algorithm : NSGA II". IEEE transactions on Evolutionary Computation, pp.182-197,2002.

[2] J. Knowles and D. Corne. The pareto archived evolution strategy: A new baseline algorithm for pareto multiobjective optimisation. Congress on Evolutionary Computation,volume 1, pages 98-105, May.ower Hotel,Washington D.C., USA,1999. IEEE Press.

[3] A. Loraschi and A. Tettamanzi. An evolutionary algorithm for portfolio selection in a downside risk framework. Working Papers in Financial Economics, 6:8-

12, June 1995.

[4]H.M.Markowitz Portfolio Selection,Journal of Finance, 1(7):77-91,1952

[5]H.M.Markowitz. Portfolio Selection: efficient diversification of investments. John Wiley & Sons, 1959.

[6] Zitzler, E., Thiele, L., Multiobjective optimization using evolutionary algorithms a comparative case study., Fifth International Conference on Parallel Problem Solving from Nature (PPSN-V),Springer, Berlin, Germany, pp. 292 - 301, 1998.

[7] E. Zitzler, E., Evolutionary algorithms for multiobjective optimization: Methods and applications, PhD thesis, Swiss Federal Institute of Technology Zurich, 1999.

[8] E. Zitzler and L. Thiele. Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach. IEEE Transactions on Evolutionary Computation, 3(4):257-271, 1999.

[9] Laura Diosan, A multi-objective evolutionary approach to the portfolio optimization problem, IEEE conference,CIMCA-IAWTIC'05.

[10] Hu, X., Shi, Y., & Eberhart, R. C. (2004). Recent advances in particle swarm. In IEEE congress on evolutionary computation 2004, Portland, Oregon, USA.