

On Pareto Set Generation in Multi-criteria Portfolio Optimization

I. Radziukynienė, A. Žilinskas

1) Vytautas Magnus University, 8 Vileikos st., Kaunas, LT-44404, Lithuania, i.radziukyniene@if.vdu.lt
2) Institute of Mathematics and Informatics, 4 Akademijos str., Vilnius, LT-08663, Lithuania, antanasz@ktl.mii.lt

Abstract – Pareto set generation methods are considered with respect to their application for multi criteria portfolio selection. Several such methods were compared experimentally including some recently proposed evolutionary methods and the method of adjustable weights. Test problems were based on standard portfolio quality criteria and data on stocks of 10 Lithuanian companies. The experimental data on the performance of the considered algorithms in different metrics are presented and discussed.

Keywords - heuristic algorithms, adjustable weights, multi-objective portfolio optimization, performance metrics

I. INTRODUCTION

Optimization methods have a long history in many financial domains. Many computational finance problems ranging from asset allocation to risk management, from option pricing to model calibration can be solved efficiently using modern optimization techniques.

The question of optimal portfolio allocation has been of long-standing interest for academics and practitioners in finance. In 1950s Harry Markowitz published his pioneering work where he has proposed a simple quadratic program for selecting a diversified portfolio of securities [1]. His model for portfolio selection can be formulated mathematically either as a problem of maximization of expected return where risk, defined as variance of return, is (upper) bounded or as a problem of minimization of risk where expected return is (lower) bounded. The classical Markowitz approach to portfolio selection reduces the problem of two criteria optimization to a one criterion optimization where the second criterion is converted to a constraint. Reduction of a multi-criteria problem to one criterion problem not always is the best method to solve multi-criteria problems especially in the case of vague comparability of criteria where whole Pareto set is of interest. Portfolio selection problem is naturally formulated as a multicriteria optimization problem where compromise between criteria crucially depends on the subjective priorities of a decision maker. Therefore Pareto set generation is an important method for portfolio selection.

Evolutionary optimization is claimed very promising for generation of Pareto sets in [2-6]. In this paper we experimentally investigate the efficiency of several evolutionary multi-criteria methods in problems of portfolio selection. For comparison we have

implemented a method based on the classical scalarization idea of weighted summation of criteria; this method has been enhanced using branch and bound like procedure for choosing weights.

Test problems were based on standard portfolio quality criteria, and data on stocks of 10 Lithuanian companies. We do not concern here in much between analytical properties of the criteria functions and such properties favourable for the considered methods; we believe, however that general (global) structure of multi-criteria portfolio selection problem will be invariant with respect to switching from criteria defined by simple analytical formula to criteria defined by complicated numerical methods.

The paper is organized as follows. In section 2 the multi-objective portfolio optimization problem is outlined, section 3 describes the selected optimization methods and their characteristics. In section 4 and 5 we discuss the used performance metrics and the experimental results. The paper is completed with the section of conclusions.

II. MULTI-OBJECTIVE PORTFOLIO OPTIMIZATION PROBLEM

Risk plays an important role in modern finance, including risk management, capital asset pricing and portfolio optimization. The problem of portfolio selection can be formulated as the problem to find an optimal strategy for allocating wealth among a number of securities (investment) and to obtain an optimal risk-return trade-off. The portfolio optimization problem may be formulated in various ways depending on the selection of the objective functions, the definition of the decision variables, and the particular constraints underlying the specific situation. Beyond the expected return and variance of return, like in Markowitz portfolio model [1], the additional objective function can include number of securities in a portfolio, turnover, amount of short selling, dividend, liquidity, excess return over of a benchmark random variable and other [7]. In the bank portfolio management, the additional criteria such as the prime rate, processing cost, expected default rate, probability of unexpected losses, quantity of the long-term and short-term can be considered [8]. For example, the multi-objective portfolio selection problem can include the following objectives [9]: (to be maximized) portfolio return, dividend, growth in sales, liquidity,

portfolio return over that of a benchmark, and (to be minimized) deviations from asset allocation percentages, number of securities in portfolio, turnover (i.e., costs of adjustment), maximum investment proportion weight, amount of short selling.

We considered two multi-objective portfolio problems. The first problem was based on a simple two objectives portfolio model including the standard deviation of the returns and mean of the returns, where the return R_i is one month return of stock i ; return means percentage change in value. The second problem included three objectives, where annual dividend yield is added to two above mentioned objectives. For the experiment we used a dataset of 10 Lithuanian companies' stock data from Lithuanian market.

There are many methods to attack multi-criteria optimization problems including those aimed at generating a set of Pareto optimal solutions. However, the problem to find the whole Pareto set normally (e.g. in a case of continuum cardinality of Pareto set) is pure theoretical since can not be solved algorithmically. An important for applications revised version of the considered problem is a problem of algorithmic construction of an appropriate approximation of Pareto set.

III. JUSTIFICATION OF SELECTION OF METHODS

In this study we investigated five methods designed to approximate Pareto set by means of a finite set of points uniformly distributed in close vicinity of the Pareto set; the terms "uniformly" and "close vicinity" are defined more precisely in the section on experimental results. The weighting method was selected due to its simple implementation and good Pareto set approximation. Three methods: Fast Pareto genetic algorithm (FastPGA) [2], Multi-Objective Cellular genetic algorithm (MOCeLL) [3], and Archive-based hybrid Scatter Search algorithm [4] were proposed during the last two years. Their efficiency for various problems has been shown in original papers, but their application to portfolio optimization problem was not yet explored. NSGA-II [5], the state-of-the-art evolutionary method, was chosen following many authors who use it as a standard for comparisons.

A. Weighting method

The weighting method is one of the most widely used methods for multi criteria optimization problems. It converts a multi criteria problem into a standard single criterion optimization problem where the objective function is a weighted sum of the criteria functions. Weights can be interpreted as indicators of the relative significance of different criteria and thus provide a solution that incorporates one's preferences. However, they can also be considered as parameters defining weights dependent solutions in the Pareto set. The

mathematical model of the weighting method takes the form of:

$$\min f(x), \quad f(x) = \sum_{i=1}^m \omega_i f_i(x), \quad (1)$$

where ω_i is the weight of i -th criterion, and $\sum_{i=1}^m \omega_i = 1$, $0 \leq \omega_i \leq 1$, $i=1, 2, \dots, m$.

Advantages. The method is relatively simple and easy to use. Assume that all $f_i(x)$ are convex functions; then for every point of Pareto set there exist weights such that this Pareto point is a minimizer of (1).

Disadvantages. The violation of the convexity assumption can imply absence of minimizers of (1) in some subsets of Pareto set. In many cases it is not an easy task to choose the set of weights defining solutions of (1) that would be well (in some sense uniformly) distributed in Pareto set. The mapping of the set of weights to the Pareto set using (1) is not necessarily a one-to-one mapping.

To generate solutions of (1) well distributed over the Pareto set, weights should be distributed in a special but a priori unknown way. We propose a branch and bound type method for iterative distribution of weights aiming at the desirable distribution of solutions of (1).

Our idea is to partition the feasible region of weights (which is a standard simplex) into sub simplices whose vertices are mapped to the Pareto set via (1). The sequential partition procedure is arranged as a branching of a tree of nodes corresponding to sub simplices. The standard simplex is accepted as the root of the tree. Branching means partition of a simplex into two sub simplices where the new vertex is the midpoint of the *favourable edge* defined later. We aim to subdivide the original simplex into sub simplices in such a way that solutions of (1) corresponding to the vertices of these sub simplices would be well distributed over the Pareto set. A simplex is eligible for branching if the longest distance between the points in the criteria space corresponding to the vertices of the considered simplex in space of weights is longer than the predefined tolerance. The *favourable edge* of the selected simplex in the space of weights is defined by the longest distance between the corresponding points in the criteria space. Branching strategy is based on depth first search. The method of adjustable weights was implemented for two and three criteria cases in MATLAB using *fminconstr* for solving minimization problem (1).

To illustrate the advantage of the proposed method of adjustable weights over the standard weighting method both methods were applied for the construction of Pareto set of two criteria portfolio optimization with the data mentioned above. Fig. 1 shows the distributions of the Pareto points found by both methods. In the standard method weights have been changed with the step for 0.05. Because of space limitation we do not describe in detail the tolerance for branching of the method of

adjustable weights; it have been set aiming at generation of similar number of points as generated by weighting method. In this experiment the number of points generated by the weighting method was equal to 21, and that generated by our method was equal to 26.

B. FastPGA

Eskandari and Geiger [2] have proposed framework named fast Pareto genetic algorithm that incorporates a new fitness assignment and solution ranking strategy for multi-objective optimization problems where each solution evaluation is relatively computationally expensive. The new ranking strategy is based on the classification of solution into two different categories according to dominance. The fitness of non-dominated solutions in the first rank is calculated by comparing each non-dominated solution with one another and assigning a fitness value computed using crowding distance. Each dominated solution in the second rank is assigned a fitness value taking into account the number of both dominating and dominated solutions. New search operators are introduced to improve the proposed method's convergence behaviour and to reduce the required computational effort. A population regulation operator is introduced to dynamically adapt the population size as needed up to a user-specified maximum population size, which is the size of the set of non-dominated solutions. FastPGA is capable of saving a significant number of solution evaluations early in the search and utilizes exploitation in a more efficient manner at later generations.

Characteristics: The regulation operator employed in FastPGA improves its performance for fast convergence, proximity to the Pareto optimal set, and solution diversity maintenance.

C. MOCeLL

Nebro *et al* [3] presented MOCeLL, a multi-objective method based on cellular model of GAs, where the concept of small neighborhood is intensively used, i.e., population member may only interact with its nearby neighbors in the breeding loop. MOCeLL uses an external archive to store the non-dominated solutions found during the execution of the method, however, the main feature characterizing MOCeLL is that a number of solutions are moved back into the population from the archive, replacing randomly selected existing population members. This is carried out with the hope of taking advantage of the search experience in order to find a Pareto set with good convergence and spread.

MOCeLL starts by creating an empty Pareto set. The Pareto set is just an additional population (the external archive) composed of a number of the non-dominated solutions found. Population members are arranged in a 2-dimensional toroidal grid, and the genetic operators are successively applied to them until the termination condition is met. Hence, for each population member, the

method consists of selecting two parents from its neighbourhood for producing an offspring.

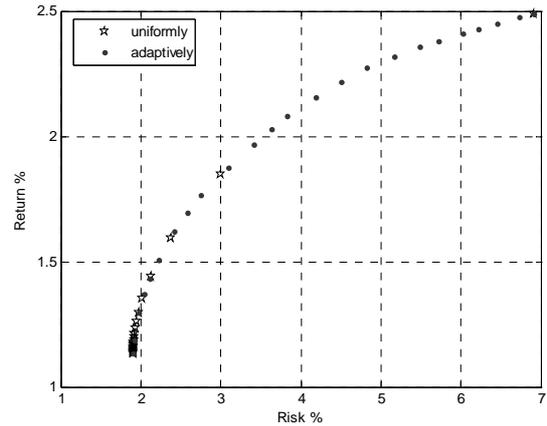


Fig. 1. Points in Pareto set generated by weighting method and by the method of adjustable weights.

An offspring is obtained applying operators of recombination and mutation. After evaluating of the offspring as the new population member it is inserted in both the auxiliary population (if it is not dominated by the current population member) and the Pareto set. Finally, after each generation, the old population is replaced by the auxiliary one, and a feedback procedure is invoked to replace a fixed number of randomly chosen population members of the population by solutions from the archive. In order to manage the insertion of solutions in the Pareto set with the goal to obtain a diverse set, a density estimator based on the crowding distance has been used. This measure is also used to remove solutions from the archive when this becomes full.

Characteristics: The method uses an external archive to store the non-dominated population members found during the search. The most salient feature of MOCeLL with respect to the other cellular approaches for multi-objective optimization is the feedback of members from archive to population.

D. AbYSS

AbYSS was introduced by Nebro *et al* [4]. It is based on the scatter search using a small population, known as the reference set, whose population members are combined to construct new solutions. Furthermore, these new population members can be improved by applying a local search method. For local search the authors proposed to use a simple (1+1) Evolution Strategy which is based on a mutation operator and a Pareto dominance test. The reference set is initialized from an initial population composed of disperse solutions, and it is updated by taking into account the solutions resulting from the local search improvement.

AbYSS combines ideas of three state-of-the-art evolutionary methods for multi criteria optimization. On the one hand, an external archive is used to store the non-

dominated solutions found during the search, following the scheme applied by PAES [5], but using the crowding distance of NSGA-II [6] as a niching measure instead of the adaptive grid used by PAES; on the other hand, the selection of solutions from the initial set to build the reference set applies the density estimation used by SPEA2 [4].

Characteristics: It uses an external archive to store the non-dominated population members found during the search. Salient features of AbYSS are the feedback of population members from the archive to the initial set in the restart phase of the scatter search, as well as the combination of two different density estimators in different parts of the search.

E. NSGA-II

The evolutionary method for multi-criteria optimization NSGA-II contains three main operators: a non-dominated sorting, density estimation, and a crowded comparison [6]. Starting from a random population the mentioned operators govern evolution whose aim is uniform covering of Pareto set.

Non-dominated sorting maintains a population of non dominated members: if a descendant is dominated, it immediately dies, otherwise it becomes a member of population; all members of parent generation who are dominated by descendants die.

The density at the particular point is measured as the average distance between the considered point and two points representing the neighbour (left and right) population members.

The crowded comparison operator defines selection for crossover oriented to increase the spread of current approximation of Pareto front. Population members are ranked taking into account “seniority” (generation number) and local crowding distance.

The worst-case complexity of NSGA-II algorithm is $O(mN^2)$, where N is the population size and m is the number of objectives [6].

Characteristics: This method is of the lower computational complexity than that of its predecessor NSGA. Elitism is maintained. No sharing parameter needs to be chosen because sharing is replaced by crowded-comparison to reduce computations.

IV. DESIGN OF EXPERIMENT

There are some publications on comparison of evolutionary methods performance according different aspects. For example, in the study presented in [11] some general conclusions are drawn, however they could not be directly applied to a specific problem.

For assessing of the considered methods several different performance measures can be taken into account: the distance between the approximated Pareto set generated by the considered method and the true Pareto set, the spread of the solutions, and computational time. To determine the first measure the true Pareto set should be known. In this investigation we didn't know

true Pareto sets. Therefore, the best approximation found by means of combining results of all considered methods was used instead of true Pareto set. Although, computational time is one of the most important performance measures in comparison of optimization methods, it has not been included in our study. This can be justified only by the severe differences in implementations of the algorithms making direct comparison of running times unfair.

We compared methods according to three performance measures:

1. *Generational distance* (GD) shows how far the approximation is from the true Pareto set [4].
2. *Inverted generational distance* (IGD) [10]. This quality indicator is used to measure how far the elements are in the Pareto optimal set from those in the set of non-dominated vectors found.
3. *Hypervolume* (HV) [4]. This quality indicator calculates the volume (in the objective space) covered by members of a non-dominated set of solutions. Methods with larger values of HV are desirable.

Before evaluating the fitness function in FastPGA, MOCeLL, AbYSS, and NSGAII the proportions of stocks in the portfolio were normalized as in reference [8]. These methods were run with different parameters that were recommended by the authors [2-4, 6]. They are given in Table I.

To evaluate each of these methods, while solving two objectives portfolio optimization problem we performed three series of experiments. First, we ran all the methods for 15000 function evaluations, and then repeated them with the execution of 25000 and 35000 function evaluations as the stopping condition. In the case of three objectives portfolio, two series of experiments with 25000 and 50000 function evaluations have been performed. For each problem we have executed 50 independent runs.

TABLE I
METHOD PARAMETERS

Method	Parameters
FastPGA	Maximum population = 100, initial population = 100, crossover probability = 1.0
MOCeLL	Population = 100, archive = 100, crossover probability = 0.9
AbYSS	Population = 20, archive = 100, crossover probability = 1.0, setref1=10, setref2=10
NSGAII	Population = 100, crossover probability = 0.9

V. DISCUSSION OF EXPERIMENTAL RESULTS

The experiments were performed to compare performance of the selected evolutionary methods. The performance criteria of interest were calculated also for the method of adjustable weights usable for scaling of results. The experimental results for the two criteria problem are given in Table II. The averages and standard deviations of all performance measures for each evolutionary method were calculated from the samples of

50 runs. It follows from the Table II that MOCeLL outperforms the other methods in all cases except of the case of maximal number of evaluations set equal to 25000, and comparison with respect to Inverted general distance; in the later case NSGAII was considerably better. The differences in performance weaken with increase of maximal number of evaluations as it is well illustrated by the results of experiments with this parameter set equal to 35000.

The upper part of the Pareto set shown in Fig.1 is most difficult to reveal for all evolutionary methods. Approximation of the Pareto set with a curve of changing curvature shows that the curve is flattening at its upper part. Similar dependence between flatness of the Pareto front and decrease of quality of its approximation using evolutionary methods is mentioned also by the other authors. For illustration of this phenomenon the points generated by all methods in the mentioned part of Pareto set are presented in Fig. 2; maximal number of function evaluations was fixed equal to 25000. It can be noticed that the method of adjustable weights does not suffer from flattening of the Pareto front.

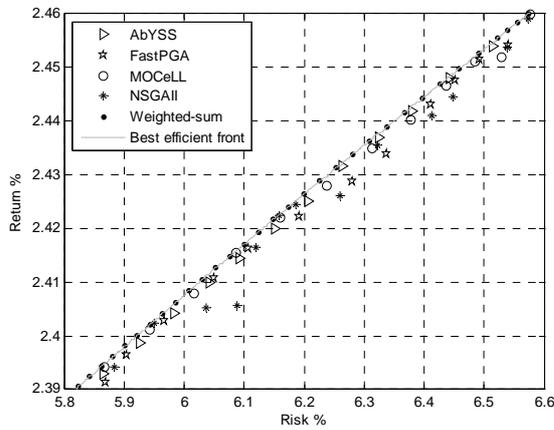


Fig. 2. Pareto sets of AbYSS, FastPGA, MOCeLL, NSGAII, and Weighted-sum method.

The experimental results of three criteria problem are given in Table III. These results show that the best of the evolutionary methods according GD and HV measures is MOCeLL. NSGAII and FastPGA are the best methods with respect to IGD. It can be noticed that GD value of weighting method is equal zero; this means that all solutions lie precisely in Pareto set.

The parts of Pareto sets obtained by evolutionary methods with 50000 function evaluations and the weighted-sum method are presented in Fig. 3. From this figure it can be seen how much the quality of approximation of Pareto set by evolutionary methods is behind of that of weighting method; this indicates the perspective for improvement of evolutionary methods.

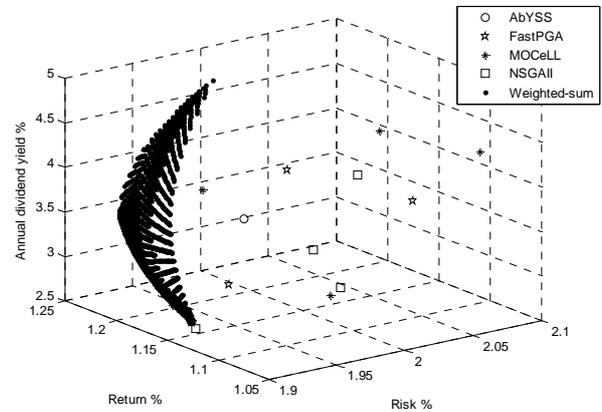


Fig. 3. Pareto sets of AbYSS, FastPGA, MOCeLL, NSGAII, and Weighting method.

TABLE II
PERFORMANCE METRICS OF TWO OBJECTIVES PORTFOLIO PROBLEM

Method	GD		IGD		HV	
	Avg.	Std.	Avg.	Std.	Avg.	Std.
Weighting	0.0		2.97 E-5		0.8861	
Maximum number of evaluations is 15000						
AbYSS	5.470E-4	4.092 E-4	1.86 E-3	2.141 E-3	0.8552	0.0331
FastPGA	4.122E-4	2.808 E-4	2.28 E-3	1.636 E-3	0.8563	0.0238
MOCeLL	3.100 E-4	2.315 E-4	1.226E-3	1.653 E-3	0.8696	0.0212
NSGAII	4.173 E-4	2.751 E-4	1.4E-3	1.238E-3	0.8694	0.0141
Maximum number of evaluations is 25000						
AbYSS	2.056 E-4	5.95 E-5	1.60 E-4	5.586 E-4	0.8821	0.0055
FastPGA	2.260 E-4	7.04 E-5	4.88 E-4	9.774 E-4	0.8786	0.0126
MOCeLL	9.289 E-5	1.88 E-5	1.11 E-4	2.918 E-4	0.8836	0.0019
NSGAII	2.356 E-4	2.87 E-5	9.9 E-5	1.28 E-5	0.8823	0.0003
Maximum number of evaluations is 35000						
AbYSS	1.631 E-4	3.57 E-5	7.2 E-5	2.4 E-6	0.8834	3.2E-4
FastPGA	2.108 E-4	2.72 E-5	1.04 E-4	1.155 E-4	0.8827	4.7E-4
MOCeLL	6.339 E-5	8.9 E-6	6.7 E-5	9 E-7	0.8843	8.5E-5
NSGAII	2.415 E-4	3.02 E-5	9.7 E-5	4.7 E-6	0.8824	1.8E-4

TABLE III
PERFORMANCE METRICS OF THREE OBJECTIVES PORTFOLIO PROBLEM

Method	GD		IGD		HV	
	Avg.	Std.	Avg.	Std.	Avg.	Std.
Weighting	0.0		1.48E-4		0.7355	
Maximum number of evaluations is 25000						
AbYSS	1.438E-3	4.68E-4	2.16E-4	1.16E-4	0.7148	0.0063
FastPGA	1.418E-3	3.85E-4	2.08E-4	1.9E-5	0.7161	0.0021
MOCeLL	1.164E-3	3.62E-4	2.12E-4	1.9E-5	0.7181	0.0013
NSGAI	1.327E-3	3.02E-4	2.10E-4	2.2E-5	0.7154	0.0019
Maximum number of evaluations is 50000						
AbYSS	1.221E-3	4.04E-4	2.126E-4	1.9 E-5	0.7163	0.0015
FastPGA	1.239E-3	3.59E-4	2.061E-4	1.5 E-5	0.7177	0.0014
MOCeLL	1.14 E-3	2.77E-4	2.119E-4	1.8 E-5	0.7194	0.0012
NSGAI	1.648E-3	5.05E-4	2.118E-4	2.2 E-5	0.7163	0.0015

VI. CONCLUSIONS

From the results of three sets of experiments for two criteria portfolio optimization it follows that MOCeLL is the best of four considered evolutionary methods with respect to all three performance criteria.

The results of two sets of experiments with these methods for three criteria portfolio optimization reveal that MOCeLL provides the best results in terms of Hypervolume, and Generational distance, but is slightly outperformed by FastPga with respect to the Inverted generational distance.

The evaluated performance criteria of evolutionary methods are only slightly worse than those of method of adjustable weights who is advantageous in the considered cases.

Summarizing the results it seems promising to investigate a possibility of combining the advantages of evolutionary methods with those of the method of adjustable weights.

REFERENCES

- [1] H. Markowitz, "Portfolio selection," *Journal of Finance*, vol. 7, pp. 77-91, 1952.
- [2] H. Eskandari and C.D. Geiger, "A Fast Pareto Genetic Algorithm approach for solving expensive multiobjective optimization problems," *Journal of Heuristics*, in press.
- [3] A.J. Nebro, J.J. Durillo, F. Luna, B. Dorronsoro, E. Alba, "A Cellular Genetic Algorithm for multiobjective optimization," *Proceedings of NICSO 2006*, pp. 25-36. Granada, Spain, June 2006.
- [4] A.J. Nebro, F. Luna, E. Alba, A. Beham, B. Dorronsoro, "AbYSS: Adapting Scatter Search for multiobjective optimization," Tech Rep. ITI-2006-2, Departamento de Lenguajes y Ciencias de la Computación, University of Málaga, 2006.
- [5] J.D. Knowles, D.W. Corne, "Approximating the nondominated front using the Pareto Archived Evolution Strategy," *Evolutionary Computation*, vol. 8, Nr. 2, 2000, pp. 149-172.
- [6] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, Nr. 2, 2002, pp. 182-197.
- [7] Ehgott, C. Waters, R. N. Gasimov, O. Ustun, "Multiobjective Programming and Multiattribute Utility Functions in Portfolio Optimization," 2006. Available on <http://www.esc.auckland.ac.nz/research/tech/esc-tr-639.pdf>
- [8] A. Mukerjee, R. Biswas, K. Deb, and A. P. Mathur. Multi-objective evolutionary algorithm for the risk-return trade-off in bank loan management. *International Transactions in Operational research*, vol. 9, 2002, pp. 583-597.
- [9] R. E. Steuer, Y. Qi, and M. Hirschberger. Portfolio Selection in the Presence of Multiple Criteria. In C. Zopounidis, M. Doumpos, P. M. Pardalos (Eds) *Handbook of Financial Engineering*, Springer, 2008.
- [10] D. A. Van Veldhuizen and G. B. Lamont, "Multiobjective Evolutionary Algorithm Research: A History and Analysis," Dept. Elec. Comput. Eng., Graduate School of Eng., Air Force Inst. Technol., Wright-Patterson, AFB, OH, Tech. Rep. TR-98-03, 1998.
- [11] E. Zitzler, K. Deb, L. Thiele, "Comparison of multiobjective evolutionary algorithms: empirical results," *Evolutionary Computation* vol. 8, Nr. 2, 2000, pp. 173-195.