A Local Search Based Evolutionary Multi-objective Optimization Approach for Fast and Accurate Convergence

Karthik Sindhya[1,3], Kalyanmoy Deb[1,3], and Kaisa Miettinen[2]

[1] Department of Mechanical Engineering
Indian Institute of Technology Kanpur, PIN 208016, India
ksindhya@iitk.ac.in, deb@iitk.ac.in
[2] Department of Mathematical Information Technology, P.O. Box 35 (Agora),
FI-40014 University of Jyväskylä, Finland
kaisa.miettinen@jyu.fi
[3] Department of Business Technology, Helsinki School of Economics, P.O. Box 1210,
FI-00101 Helsinki, Finland

Abstract. A local search method is often introduced in an evolutionary optimization technique to enhance its speed and accuracy of convergence to true optimal solutions. In multi-objective optimization problems, the implementation of a local search is a non-trivial task, as determining a goal for the local search in presence of multiple conflicting objectives becomes a difficult proposition. In this paper, we borrow a multiple criteria decision making concept of employing a reference point based approach of minimizing an achievement scalarizing function and include it as a search operator of an EMO algorithm. Simulation results with NSGA-II on a number of two to four-objective problems with and without the local search approach clearly show the importance of local search in aiding a computationally faster and more accurate convergence to Pareto-optimal solutions. The concept is now ready to be coupled with a faster and more accurate diversity-preserving procedure to make the overall procedure a competitive algorithm for multi-objective optimization.

1 Introduction

Evolutionary multi-objective optimization (EMO) algorithms are often criticized for their lack of a theoretical convergence proof to the true Pareto-optimal front. Although theoretical time complexity estimates of certain specific EMO algorithms exist [15] in solving specific test problems, in most problems a proof of convergence with a finite computational effort is missing. However, EMO algorithms also lack a theoretical proof of convergence to even on a local Pareto-optimal front. Moreover, a past study [14] has demonstrated and argued that EMO algorithm with a finite size archive for storing non-dominated solutions may allow an evolving population to fluctuate (convergence to the Pareto-optimal front followed by a departure of some solutions out of the front). This phenomenon can happen due to constant emphasis of diversity maintenance

G. Rudolph et al. (Eds.): PPSN X, LNCS 5199, pp. 815–824, 2008.

pursued in EMO. To make the diversity among obtained non-dominated solutions better, a Pareto-optimal solution may be sacrificed to accept a non-Pareto-optimal solution.

Despite the lack of theoretical convergence properties of EMO algorithms, they are increasingly being used in many applications, due to reasons such as, user's satisfaction with a near Pareto-optimal solution, difficulty in implementation of a local search procedure in a multi-objective context and in discrete and combinatorial optimization problems, optimality of a solution is impossible to verify. There exist many other practical optimization problems for which solutions close to the true Pareto-optimal front are desired with as low a computational effort as possible.

The use of local search in EMO has enjoyed a lot of attention in recent past, to make EMO algorithms converge faster on to the true Pareto-optimal front [3]. Here, we briefly mention some representative studies. A hybrid algorithm (S-MOGLS) using weighted sum of multiple objectives as fitness function was proposed in [9]. A neighborhood search (NS) was used as a local search which was then applied to all offspring solutions generated by NSGA-II. The algorithm C-MOGLS developed in [17] combined cellular multi-objective genetic algorithm and NS as a local search. The local search in this approach was applied using weighted sum of multiple objectives as fitness function to all non-dominated solutions in each generation. A new genetic local search algorithm which uses hybridization of recombination operators with a neighborhood based local search was presented in [10]. A random utility function was optimized locally in this algorithm. The local search was applied on all offspring solutions. In [7], a weighted sum of multiple objectives as fitness function was used and two approaches were presented hybridizing NSGA-II, a posteriori approach in which the local search is applied on all non-dominated solutions obtained after the NSGA-II simulation and an online approach in which the local search is applied to all offspring generated in each generation of NSGA-II. M-PAES which is a population version of multi-objective evolution strategy (PAES) was proposed in [11]. The search is enhanced by the use of (1+1)-ES as a local search. Furthermore, a new hybrid algorithm which uses Pareto descent method (PDM) as local search method was proposed in [8]. PDM finds feasible Pareto descent directions by solving computationally inexpensive linear programming problems.

Based on these studies and others from the literature, we observe two main approaches of using a local search with EMO. First, most studies implement local search as a refinement of solutions found by EMO. Second, when implemented within EMO, a local search is usually applied to all offspring solutions with a naive neighborhood based procedure. Not much effort has been given in borrowing more effective multiple criteria decision making (MCDM) [1] ideas in local search. In this paper, we propose the idea of one such hybrid approach in which EMO and local search are coupled and the latter is used to solve an augmented achievement scalarizing function (ASF). Since the solution to an augmented ASF is always a properly Pareto-optimal solution, the overall hybrid approach is shown to have a better convergence property than the EMO procedure alone.

The rest of this paper is organized as follows. In Section 2, we discuss the motivation of using a local search as a part of multi-objective optimization. The proposed hybrid approach is described in Section 3. Then, Section 4 describes results obtained by the hybrid approach and compared with original NSGA-II on test problems and briefly discusses possible approaches to ensure diversity in a hybrid approach. Finally, conclusions are drawn in Section 5.

2 Local Search in EMO

A local search is usually applied to improve the solution(s) obtained by an approximate optimizer. Here, we discuss the motivation for using a local search in a complex optimization task.

Optimization in practice involves objective function landscapes which may be multimodal, nondifferentiable, discrete, and involve many other complexities. It is unrealistic to expect a single optimization algorithm to be computationally efficient in handling different vagaries of function landscapes. A combination of two types of algorithms – an approximate global optimizer and an accurate local search – is one way to tackle the problem. The global optimizer searches the entire landscape to find the most promising region(s) with multiple points, while the local search begins its search from a particular solution and converges to a locally optimal solution. Thus, the roles of global and local optimizers are used to negotiate different function landscapes. Both optimization tasks are important and a balance of the extent of their searches is necessary for the overall procedure to converge to the true globally optimal solution.

In this study, we consider a population-based evolutionary algorithm as a global optimizer and a (gradient-based) mathematical programming method as a local search procedure, as they fit well with the above description of local and global optimizers. There are at least two different ways they can be hybridized: a *serial* and a *concurrent* approach. In a serial approach, global and local searches are applied serially one after the other with appropriate termination conditions. This switchover to local search from the global optimizer is not easy to fix a priori on any unknown problem, as a delay in terminating global solver can consume excess function evaluations and an early termination shall yield a locally optimal solution. For terminating a local search, a standard procedure such as the error in violations of Karush-Kuhn-Tucker (KKT) [12] conditions of optimality to be within a limit can be used. In the concurrent approach, local search is embedded within a global optimizer so that some or all intermediate solutions are modified by the local search. For example, in the EMO framework, the local search may be considered as an additional EMO operator which attempts to bring an intermediate solution to a locally Pareto-optimal solution. The termination criterion of the local search can be a standard one (as described above), but the termination of the global optimizer need not be ad-hoc, as in the serial approach, but can be based on whether there is an improvement in the locally optimal solutions over past few iterations. Due to the above advantages, we restrict ourselves to the concurrent approach of hybridization in this study.

In the case of solving single-objective optimization problems, the objective function to be optimized in the local search can be the same as that used in the global optimizer, as the main goal in this task is to find the global optimum of a single function landscape. Figure 1 illustrates this aspect. However, in handling multiple conflicting objectives, a local search faces an additional difficulty of choosing an appropriate single objective for its search. Since multiple conflicting objectives are of interest here, it is not fair to choose one particular objective function among the conflicting ones for the local search.

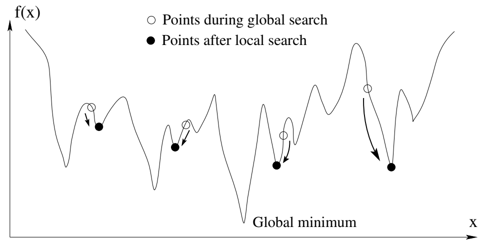

Fig. 1. Local and global searches can use the same objective function in a single-objective optimization

Thus, we realize that an implementation of a local search is nontrivial yet important in the context of multi-objective optimization. Despite, the existence of many local search approaches described in Section 1, a directed and computationally faster optimization approach is necessary for the local search.

In the following section, we suggest a widely used reference point based approach [18] from the MCDM field to be used in local search.

3 Proposed Local Search Based EMO

We propose a hybrid approach where we use the NSGA-II method [5] as the EMO procedure and hybridize it with an ASF (based on a reference point) which is solved with a local search method.

In the t-th generation of the NSGA-II procedure, an offspring population Q_t is created by using selection, recombination and mutation operators from the parent population P_t. Thereafter, each member of Q_t is evaluated and checked with a probability p_l for its improvement with the local search procedure. After the local search operations are performed, parent and offspring populations are combined together and a non-dominated sorting is performed. Thereafter, the NSGA-II procedure continues as usual.

The local search is started from an offspring solution $\mathbf{y}$ (having objective vector $\mathbf{f}(\mathbf{y})$). The local search procedure minimizes the following augmented achievement scalarizing function [18]:

$$\begin{aligned} &\text{minimize}_{\mathbf{x}} \max_{i=1}^{M} \frac{f_i(\mathbf{x})-z_i}{f_i^{\max}-f_i^{\min}} + \rho \sum_{j=1}^{M} \frac{f_j(\mathbf{x})-z_j}{f_j^{\max}-f_j^{\min}}, \\ &\text{subject to } \mathbf{x} \in \mathcal{S}, \end{aligned} \tag{1}$$

where $\mathcal{S}$ is the feasible decision variable space, $\mathbf{z}=\mathbf{f}(\mathbf{y})$ is the so-called reference point, and $f_i^{\max}$ and $f_i^{\min}$ are the maximum and minimum objective values of the parent population P_t used for normalization.

By solving (1) we project the reference point (i.e. the solution produced by the EMO algorithm) onto the Pareto optimal front. The second term in the objective function (1) ensures that the local search will converge to a properly Pareto-optimal solution [16,18]. A small value of $\rho = 10^{-2}$ is used in this study. This local search procedure allows a directed search dictated by the $w_i = 1/(f_i^{\max} - f_i^{\min})$ term, as shown in Figure 2. The local search is terminated if any of the following conditions is met: (i) a maximum of 25 iterations is elapsed (here, maximum iterations are fixed to prevent excessive function evaluations during initial stages of algorithm) (ii) a KKT error value of 0.001 is achieved, or (iii) a maximum difference of 10^{-6} in any variable in two successive iterations is achieved.

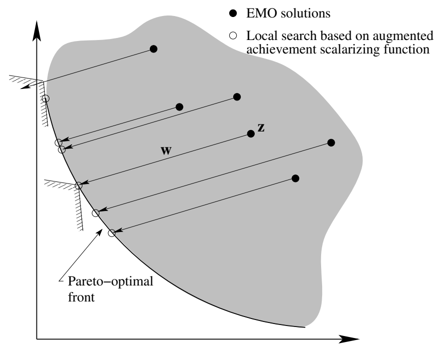

Fig. 2. Proposed local search based on augmented achievement scalarizing function

In this study, we use a probability of local search p_l which periodically increases and drops linearly with generations. Starting from zero at the initial generation, the probability rises to 0.01 in $(0.5N - 1)$ generations (where N is the population size) and drops to zero in $t = 0.5N$ generations. This means that, when $N = 100$, and generation $= (0.5N - 1)$, on an average one solution in the entire population gets modified by the local search. The initial generations have a smaller local search probability, as typically the population is far from the Pareto-optimal front and the local search may mostly produce extreme Pareto-optimal solutions. The probability increases linearly as more solutions may need to be modified using the local search procedure to ensure convergence to the Pareto-optimal front. p_l goes to zero after each period to prevent loss in diversity both during these initial phases and when the population approaches the Pareto-optimal front.

To terminate the hybrid approach, the normal stopping criteria of EMO such as fixed maximum number of generations can be applied. Alternatively, it is time to stop when the local search produces no significant change. Then the local search should be applied to the entire final population. Here, for testing purposes we use a stopping criterion based on the discrepancy (we call it an 'error metric') in objective f_M between obtained solution and corresponding f_M value obtained by substituting other objective values (f_1 to f_{M-1}) in the Pareto-optimal relationship ($f_M = f_M(f_1, \ldots, f_{M-1})$) is calculated for each current non-dominated solutions. If the sum of the square of errors generated by all non-dominated solutions is less than or equal to 0.001, the hybrid approach is terminated. This termination criterion ensures that all obtained solutions are close to the true Pareto-optimal front. It must be noted, that this termination criterion is used only to test the efficacy of our algorithm and cannot be applied to any general problem. To, check the diversity of obtained set of non-dominated solutions, we calculate the hypervolume measure (HV) of the obtained set of points [4]. Thereafter, the normalized difference in hypervolume measure (NDHV = (HV*-HV)/HV*, where HV* is the hypervolume of the true Pareto-optimal front) is computed and compared with the same obtained for the solutions of the original NSGA-II.

Although not obvious, we argue here that the use of a local search procedure within an EMO shall constitute a computationally faster approach than without the use of a local search to EMO procedure. The occasional use of the local search procedure will introduce a few elite solutions in the population. Under EMO operators, these solutions will then get an opportunity to recombine with other population members and exchange variables between them, which may cause more non-Pareto-optimal solutions to come closer to the Pareto-optimal front. This hybrid approach allows the population to converge faster near the Pareto-optimal front.

4 Results and Discussion

We apply the proposed hybrid approach on a number of two to four objective test problems. As a local solver we use SQP from KNITRO [2]. To compare the speed and accuracy of our hybrid approach with the original NSGA-II, both algorithms are terminated when the average error metric value is smaller than 0.001 and the number of function calls needed in each case are recorded. For each problem, we also compute and compare the NDHV.

For this study, we consider four bi-objective test problems (ZDT1, ZDT2, ZDT3 and ZDT4) and two three-objective test problems (DTLZ1 and DTLZ2) and one four-objective problem (DTLZ2). For bi-objective problems, we have 100 population members and for three and four objectives, we have used 200 population members. Crossover probability of 0.9, SBX distribution index [4] of 15, mutation probability of 0.1^{M-1} (reduced probability with number of objectives due to increased maintenance of diversity by crowding distance operator), and mutation distribution index of 20 are used. Table 1 show the best, median

Table 1. Comparison of the number of function calls for the hybrid approach and original NSGA-II. Algorithms are terminated when a fixed level of convergence is achieved.

Test	Original NSGA-II			Hybrid approach		
Problem	Best	Median	Worst	Best	Median	Worst
ZDT1	19,400	20,900	24,600	**4,665**	7,554	8,580
ZDT2	20,600	21,700	23,200	**4,826**	6,351	7,198
ZDT3	21,900	23,500	26,300	**10,736**	16,731	22,137
ZDT4	27,100	34,500	60,300	**6,003**	7,658	14,479
3-DTLZ1	59,800	76,200	97,200	**48,352**	60,323	76,890
3-DTLZ2	38,200	54,000	93,800	**32,611**	61,508	69,650
4-DTLZ2	52,200	68,200	120,800	**45,005**	61,879	114,899

Table 2. Comparison of NDHV for the hybrid approach and original NSGA-II. (smaller value is better).

Test Problem	Original NSGA-II			Hybrid approach		
	Best	Median	Worst	Best	Median	Worst
ZDT1	0.0043	0.0047	0.0054	**0.0034**	0.0042	0.1630
ZDT2	0.0044	0.0053	0.0064	**0.0037**	0.0070	0.0499
ZDT3	0.0012	0.0016	0.0023	**0.0007**	0.0009	0.0010
ZDT4	0.0042	0.0047	0.0055	**0.0037**	0.0106	0.223
3-DTLZ1	0.0196	0.0341	0.0403	**0.0187**	0.0224	0.0296
3-DTLZ2	0.0000	0.0000	0.0000	0.0000	0.0001	0.0030
4-DTLZ2	0.0049	0.0066	0.0086	**0.0048**	0.0067	0.0085

and worst function calls for 10 runs started from 10 identical initial populations for both the hybrid approach and the original NSGA-II procedure. The better algorithm in terms of the smallest number of required function calls is marked in bold face. It is clear that for bi-objective problems the convergence is much faster with the hybrid approach. For more objectives, better results are observed with the hybrid approach, but the difference in the performance seems to reduce with an increase in the number of objectives. We suspect that this behavior is due to the degraded performance of domination-based EMO approaches with an increased number of objectives [4]. Although the hybrid approach did not explicitly introduce a mechanism to maintain diversity except NSGA-II's crowding distance operator, we present NDHV values in Table 2. It is interesting that in most cases the proposed hybrid approach is able to find a well-distributed set of converged points. In the case of ZDT4, relatively higher median and worst values, together with a prescribed error measure of 0.001 and a smaller number of function calls, indicate desired convergence at a faster rate, but at the expense of needed diversity in some simulation runs. Since in most problems, a function evaluation is most time consuming, we stress here on number of function calls, rather than exact computational time.

Other ideas could be considered for ensuring diversity, such as: Firstly, the local search direction can be biased differently for different EMO solutions based on the location of EMO solution and on the undiscovered regions of the Pareto-optimal front. Secondly, the crowding distance operator of NSGA-II can be replaced with a better diversity preserving procedure, such as clustering [13,19] and lastly, instead of using a generational evolutionary optimization approach such as NSGA-II, a steady-state procedure (such as in epsilon-MOEA [6]) may be adopted. In this way, every new solution created by the hybrid approach can be evaluated for its convergence and diversity enhancement properties to the rest of the non-dominated solutions of the EMO population.

We now discuss the working principle of the local search by examining its performance on the ZDT1 and ZDT2 problems. In Figures 3 and 4, we plot the average error metric values versus generation counter for both the approaches. The vertical lines indicate the generations at which at least one local search is executed. It is interesting to observe that soon after the first local search has

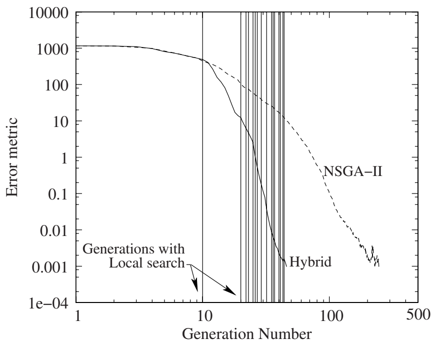

Fig. 3. Average error metric for ZDT1

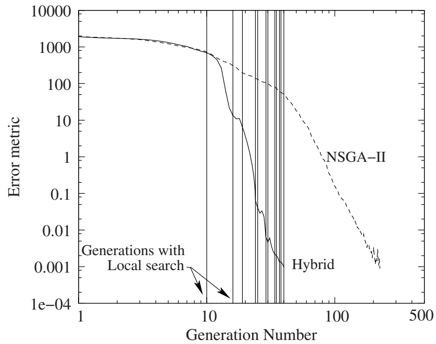

Fig. 4. Average error metric for ZDT2

taken place at generation 10, the population gathered momentum to move towards the Pareto-optimal frontier for both problems. Figure 5 shows all 100 population members at the start of generations 10, 11, 13 and 15. The first local search takes place at generation 10 and a weak Pareto-optimal solution on f_2 axis (marked with a circle) was found (25 iterations of local search were not enough to find proper Pareto optimal solution). The presence of this solution in the population and its subsequent recombination with other population members caused them to come closer to the Pareto-optimal front, thereby providing the speed of convergence. Starting from the same initial population, the original NSGA-II, although maintained a similar error metric value till generation 10, failed to keep pace with the hybrid approach thereafter. Figure 6 shows the populations at the above generations to demonstrate the slow nature of convergence of the original NSGA-II. A similar phenomenon is observed with test problem ZDT2 (Figure 4).

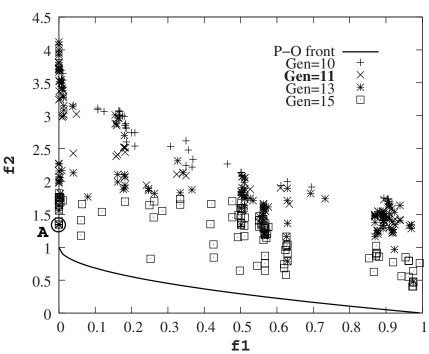

Fig. 5. Populations approach the Pareto-optimal front faster in the hybrid approach

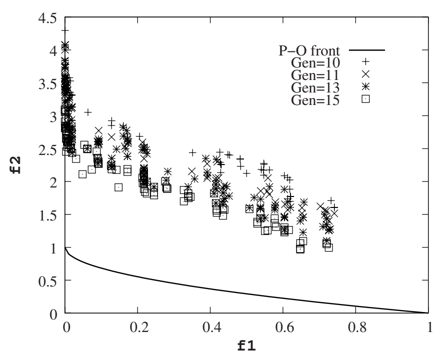

Fig. 6. Populations approach the Pareto-optimal front slowly in the original NSGA-II approach

5 Conclusions

In this paper, we have argued that an efficient implementation of a local search procedure in an EMO algorithm is not straightforward. To take advantage of fast and accurate convergence to Pareto-optimal solutions, EMO algorithms must use a directed and provable local search procedure. In this study, we have suggested the use of an augmented achievement scalarizing function to be solved with a local search method. The local search procedure has been implemented as an additional operator and applied to EMO populations with a varying probability. On a number of standard test problems involving two to four objectives, we have observed that our proposed hybrid approach with NSGA-II is computationally faster than the original NSGA-II procedure in finding solutions which are close to the Pareto-optimal front.

Achieving convergence of solutions to the Pareto front is just half of our quest, as diversity of obtained solutions is also vital. Although in this study we rely on the NSGA-II crowding distance operator for diversity preservation, the faster and accurate convergence achieved with the proposed local search is now ready to be coupled with a more efficient diversity ensuring operator (currently under study). However the present study has clearly shown the advantage and potential of hybridizing local search in a fast and accurate computation of Pareto-optimal solutions.

Acknowledgements

The support by the Academy of Finland (grant 118319), the Foundation of Helsinki School Economics and the Jenny and Antti Wihuri Foundation (Finland) is highly appreciated.

References

1. Belton, V., Stewart, T.: Multiple Criteria Decision Analysis: An integrated Approach. Kluwer, Dordrecht (2002)
2. Byrd, R.H., Nocedal, J., Waltz, R.A.: KNITRO: An integrated package for nonlinear optimization. In: di Pillo, G., Roma, M. (eds.) Large-Scale Nonlinear Optimization, pp. 35–59. Springer, Heidelberg (2006)
3. Coello Coello, C.A., VanVeldhuizen, D.A., Lamont, G.: Evolutionary Algorithms for Solving Multi-Objective Problems, 2nd edn. Springer, New York (2007)
4. Deb, K.: Multi-objective Optimization using Evolutionary Algorithms. Wiley, Chichester (2001)
5. Deb, K., Agrawal, S., Pratap, A., Meyarivan, T.: A fast and elitist multi-objective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6(2), 182–197 (2002)
6. Deb, K., Mohan, M., Mishra, S.: Towards a quick computation of well-spread Pareto-optimal solutions. In: Fonseca, C.M., Fleming, P.J., Zitzler, E., Deb, K., Thiele, L. (eds.) EMO 2003. LNCS, vol. 2632, pp. 222–236. Springer, Heidelberg (2003)

7. Goel, T., Deb, K.: Hybrid methods for multi-objective evolutionary algorithms. In: Proceedings of the Fourth Asia-Pacific Conference on Simulated Evolution and Learning (SEAL 2002), pp. 188–192 (2002)
8. Harada, K., Sakuma, J., Kobayashi, S.: Local search for multiobjective function optimization: Pareto descent method. In: Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2006), pp. 659–666 (2006)
9. Ishibuchi, H., Narukawa, K.: Some issues on the implementation of local search in evolutionary multiobjective optimization. In: Deb, K. (ed.) GECCO 2004. LNCS, vol. 3102, pp. 1246–1258. Springer, Heidelberg (2004)
10. Jaskiewicz, A.: Genetic local search for multiple objective combinatorial optimization. European Journal of Operational Research 371(1), 50–71 (2002)
11. Knowles, J.D., Corne, D.: M-PAES: A memetic algorithm for multiobjective optimization. In: Proceedings of Congress on Evolutionary Computation (CEC 2000), pp. 325–332 (2000)
12. Kuhn, H.W., Tucker, A.W.: Nonlinear Programming. In: Neyman, J. (ed.) Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, pp. 481–492. University of California Press, Berkeley (1951)
13. Kukkonen, S., Deb, K.: A fast and effective method for pruning of non-dominated solutions in many-objective problems. In: Runarsson, T.P., Beyer, H.-G., Burke, E.K., Merelo-Guervós, J.J., Whitley, L.D., Yao, X. (eds.) PPSN IX 2006. LNCS, vol. 4193, pp. 553–562. Springer, Heidelberg (2006)
14. Laumanns, M., Thiele, L., Deb, K., Zitzler, E.: Combining convergence and diversity in evolutionary multi-objective optimization. Evolutionary Computation 10(3), 263–282 (2002)
15. Laumanns, M., Thiele, L., Zitzler, E., Welzl, E., Deb, K.: Running time analysis of multi-objective evolutionary algorithms on a simple discrete optimization problem. In: Guervós, J.J.M., Adamidis, P.A., Beyer, H.-G., Fernández-Villacañas, J.-L., Schwefel, H.-P. (eds.) PPSN 2002. LNCS, vol. 2439, pp. 44–53. Springer, Heidelberg (2002)
16. Miettinen, K.: Nonlinear Multiobjective Optimization. Kluwer, Boston (1999)
17. Murata, T., Nozawa, H., Tsujimura, Y., Gen, M., Ishinuchi, H.: Effect of local search on the performance of celluar multi-objective genetic algorithms for designing fuzzy rule based classification systems. In: Proceeding of the Congress on Evolutionary Computation (CEC 2002), pp. 663–668 (2002)
18. Wierzbicki, A.P.: The use of reference objectives in multiobjective optimization. In: Fandel, G., Gal, T. (eds.) Multiple Criteria Decision Making Theory and Applications, pp. 468–486. Springer, Berlin (1980)
19. Zitzler, E., Laumanns, M., Thiele, L.: SPEA2: Improving the strength Pareto evolutionary algorithm for multiobjective optimization. In: Giannakoglou, K.C., Tsahalis, D.T., Périaux, J., Papailiou, K.D., Fogarty, T. (eds.) Evolutionary Methods for Design Optimization and Control with Applications to Industrial Problems, International Center for Numerical Methods in Engineering (CIMNE), pp. 95–100 (2001)