MNK 地形上での効果的な多数目的最適化を目的とする ϵ ランキング

アギレ・エルナン[1]　田中　清[2]

この論文では，多数目的最適化における性能向上を目的に，多目的進化型アルゴリズムの選択を強化する一方法を提案する．提案法では，パレート支配によって解にランクを与えた後，解のランキングをより精緻に行うために ϵ 支配を用いたランダムなサンプリング操作を行う．サンプリング操作は，当初等しくランクされた解から選択上の優越を与えるようにサブ集合を選び，従来のパレート支配よりも広い支配領域に基づき，サンプルの良好な分布を嗜好する．我々は，提案法により NSGA-II を強化し，10 目的までの MNK 地形による非線形問題の広い範囲においてその性能を検証する．実験結果より，$3 \leq M \leq 10$ 目的の問題において，見出された解の収束性と多様性が著しく向上することが示される．

ϵ-Ranking for Effective Many Objective Optimization on MNK-Landscapes

Hernán Aguirre[3] and Kiyoshi Tanaka[4]

This work proposes a method to enhance selection of multiobjective evolutionary algorithms aiming to improve their performance on *many* objective optimization problems. The proposed method uses a randomized sampling procedure combined with ϵ-dominance to fine grain the ranking of solutions after they have been ranked by Pareto dominance. The sampling procedure chooses a subset of initially equal ranked solutions to give them selective advantage, favoring a good distribution of the sample based on dominance regions wider than conventional Pareto dominance. We enhance NSGA-II with the proposed method and test its performance on a wide range of non-linear problems using MNK-Landscapes with up to $M = 10$ objectives. Experimental results show that convergence and diversity of the solutions found can improve remarkably on $3 \leq M \leq 10$ objective problems.

1. Introduction

Multiobjective evolutionary algorithms (MOEAs)[1] optimize simultaneously two or more objective functions, aiming to find a set of compromised Pareto optimal solutions in a single run of the algorithm. Recently, there is a growing interest on applying MOEAs to solve *many* objectives optimization problems, where the number of objectives to optimize is more than three. However, state of the art MOEAs that use Pareto dominance within the selection procedure of the algorithm do not scale-up well on *many* objective problems. An important reason for this is that Pareto selection loses its discriminatory power by increasing the dimensionality of the objective space, severely deteriorating the performance of MOEAs [2].

In this work, we propose a method to enhance selection of MOEAs aiming to improve their performance on *many* objective optimization problems. The proposed method uses a randomized sampling procedure combined with ϵ-dominance [3] to fine grain the ranking of solutions after they have been ranked by Pareto dominance. The sampling procedure chooses a subset of initially equal ranked solutions to give them selective advantage, favoring a good distribution of the sample based on dominance regions wider than conventional Pareto dominance. Thus, the proposed ranking method increases selection probabilities of some of the solutions, while trying to keep a uniform search effort towards the different zones of objective space represented in the actual population.

We enhance NSGA-II [1] with the proposed method and test its performance on a broad range of subclasses of combinatorial non-linear problems, using MNK-Landscapes [2] with $2 \leq M \leq 10$ objectives and $0 \leq K \leq 50$ epistatic interactions per bit. Experimental results show that convergence and diversity of the solutions found can improve remarkably on $3 \leq M \leq 10$ objectives for all K by properly setting the parameter ϵ that determines the domination region of the sampled solutions.

2. Definitions

Let us consider a maximization multiobjective problem with M objectives:

$$\boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \cdots, f_M(\boldsymbol{x})) \qquad (1)$$

where $\boldsymbol{x} \in \mathcal{S}$ is a solution vector in the feasible solution space $\mathcal{S}$, and $f_1(\cdot), f_2(\cdot), \cdots, f_M(\cdot)$ the M objectives to be maximized. Pareto dominance and ϵ-dominance [3] are defined as follows.

[1] 信州大学ファイバーナノテク国際若手研究者育成拠点
[2] 信州大学工学部
[3] Shinshu University, Fiber-Nanotech Young Researcher Empowerment Program
[4] Shinshu University, Faculty of Engineering

Pareto dominance. A solution $\boldsymbol{x}$ is said to Pareto dominate other solution $\boldsymbol{y}$ if the two following conditions are satisfied:

$$\begin{array}{l}\forall m \in \{1, \cdots, M\}\ f_m(\boldsymbol{x}) \geq f_m(\boldsymbol{y}) \quad \wedge \\ \exists m \in \{1, \cdots, M\}\ f_m(\boldsymbol{x}) > f_m(\boldsymbol{y}).\end{array} \tag{2}$$

Here, $\boldsymbol{x}$ dominates $\boldsymbol{y}$ is denoted by $\boldsymbol{f}(\boldsymbol{x}) \succeq \boldsymbol{f}(\boldsymbol{y})$.

ϵ-dominance. A solution $\boldsymbol{x}$ is said to ϵ-dominate other solution $\boldsymbol{y}$ if the two following conditions are satisfied:

$$\begin{array}{l}\forall m \in \{1, \cdots, M\}\ (1+\epsilon) f_m(\boldsymbol{x}) \geq f_m(\boldsymbol{y}) \quad \wedge \\ \exists m \in \{1, \cdots, M\}\ (1+\epsilon) f_m(\boldsymbol{x}) > f_m(\boldsymbol{y}).\end{array} \tag{3}$$

where $\epsilon > 0.0$. Here, $\boldsymbol{x}$ ϵ-dominates $\boldsymbol{y}$ is denoted by $\boldsymbol{f}(\boldsymbol{x}) \succeq^{\epsilon} \boldsymbol{f}(\boldsymbol{y})$.

3. Method

The core of the proposed method is a randomized ϵ-sampling procedure. In the following, we first explain ϵ-sampling and then ϵ-ranking to fine grain ranking of solutions.

3.1 ϵ-sampling

ϵ-sampling assumes that there is a set of equally ranked solutions from which a subset of them should be chosen to be given selective advantage in order to proceed further with the evolutionary search. Hence, the sampling heuristic must reflect criteria that favor an effective search. Here, the sample of solutions to be given selective advantage are obtained with the following criteria,

- Extreme solutions are always part of the sample.
- Each (not extreme) sampled solution is the sole sampled representative of its area of influence. The area of influence of the sampled solutions is determined by a domination region wider than Pareto dominance, i.e. ϵ-dominance.
- Sampling of (not extreme) solutions follows a random schedule.

The first criterion tries to push the search towards the optimum values of each fitness function, aiming to find non-dominated solutions in a wide area of objective space. The second criterion assures that only one solution in a given zone of objective space is given higher rank, trying to distribute the search effort more or less uniformly among the different zones represented in the actual population. The third criterion dynamically establishes the zones that are represented in the sample. Also, in the case that there are several solutions within each zone, it increases the likelihood that the sampled solutions that will be given higher rank are different from one generation to the next, increasing the possibility of exploring wider areas of objective and variable space.

Procedure 1 illustrates the algorithm of the proposed ϵ-sampling method. Let us denote $\mathcal{A}$ the set of solutions that have been assigned the same rank based on conventional Pareto dominance. ϵ-sampling returns the sampled solutions $\mathcal{S} \subset \mathcal{A}$ that will be given selective advantage as well as the set of solutions $\mathcal{D}^{\epsilon}$ to be demoted. See that extreme solutions are the first to be assigned to the sample $\mathcal{S}$ (lines 1,2). Then, one by one, solutions are randomly chosen and included in $\mathcal{S}$ (lines 4-6), whereas solutions that lie in the wider domination region of the randomly picked solution are assigned to $\mathcal{D}^{\epsilon}$ (lines 7,8). Note that solutions in $\mathcal{D}^{\epsilon}$ are ϵ-dominated by solutions in $\mathcal{S}$, i.e. parameter ϵ is used to virtually expand the dominance area of the sampled solutions in order not to include closely located solutions in the sample.

Procedure 1 ϵ-sampling (ϵ, $\mathcal{A}$, $\mathcal{S}$, $\mathcal{D}^{\epsilon}$)

Input: ϵ-dominance factor ϵ and a set of equal ranked solutions $\mathcal{A}$

Output: $\mathcal{S}$ and $\mathcal{D}^{\epsilon}$ ($\mathcal{S} \cup \mathcal{D}^{\epsilon} = \mathcal{A}$). $\mathcal{S}$ contains extreme and ϵ-non-dominated solutions, whereas $\mathcal{D}^{\epsilon}$ contains ϵ-dominated solutions

1: $\mathcal{X} \leftarrow \{x \in \mathcal{A} \mid f_m(x) = \max(f_m(\cdot)),\ m = 1, 2, \cdots, M\}$
2: $\mathcal{S} \leftarrow \mathcal{X}$, $\mathcal{A} \leftarrow \mathcal{A} \setminus \mathcal{X}$, $\mathcal{D}^{\epsilon} \leftarrow \emptyset$
3: **while** $\mathcal{A} \neq \emptyset$ **do**
4: $\quad r \leftarrow rand(),\ 1 \leq r \leq |\mathcal{A}|$
5: $\quad z \leftarrow r$-th solution $\in \mathcal{A}$
6: $\quad \mathcal{S} \leftarrow \mathcal{S} \cup \{z\}$
7: $\quad \mathcal{Y} \leftarrow \{y \in \mathcal{A} \mid z \succeq^{\epsilon} y, z \neq y\}$
8: $\quad \mathcal{D}^{\epsilon} \leftarrow \mathcal{D}^{\epsilon} \cup \mathcal{Y}$
9: $\quad \mathcal{A} \leftarrow \mathcal{A} \setminus \{\{z\} \cup \mathcal{Y}\}$
10: **end while**
11: **return**

3.2 ϵ-ranking

ϵ-sampling works on a set of equal ranked solutions. However, within a population there could be several sets of such solutions, each with a different rank. ϵ-ranking re-ranks all sets of equal ranked solutions using ϵ-sampling. In NSGA-II, a non-domination sorting procedure is applied to the joined population of parents and offspring to classify solutions in fronts of non-dominated solutions $\mathcal{F}_i$ $(i = 1, \cdots, N_F)$. Solutions in the same front are assigned the same rank, equal to the front number they belong to.

ϵ-ranking in NSGA-II is applied at each generation after non-domination sorting to reclassify the fronts $\mathcal{F}_i$ $(i = 1, \cdots, N_F)$ into $\mathcal{F}^{\epsilon} = \{\mathcal{F}_j^{\epsilon}\}$ $(i = 1, 2, \cdots, N_F^{\epsilon})$, where $N_F^{\epsilon} \geq N_F$. **Procedure 2** describes the ϵ-ranking method. See that the reclassified front $\mathcal{F}_j^{\epsilon}$ $(j = 1, \cdots, N_F^{\epsilon})$ now contains only the sample of solutions $\mathcal{S} \subset \mathcal{F}_i$ found by ϵ-sampling (lines 9,10). Also, see that solutions $\mathcal{D}^{\epsilon}$, which are not part of the sample (line 9) are demoted by joining them with solutions of an inferior front in the next iteration of the loop (line 4). Thus, $\mathcal{F}_1^\epsilon$ contains some of the solutions initially ranked first, but $\mathcal{F}_j^\epsilon$, $j > 1$, can contain solutions that initially were ranked in different fronts.

Procedure 2 ϵ-ranking (ϵ, $\mathcal{F}$, $\mathcal{F}^\epsilon$)

Input: ϵ-dominance factor ϵ and solutions $\mathcal{F}$ classified in fronts $\mathcal{F}_i$ ($i = 1, \cdots, N_F$) by non-domination sorting

Output: $\mathcal{F}^\epsilon$, solutions re-classified in fronts $\mathcal{F}_j^\epsilon$ ($j = 1, \cdots, N_F^\epsilon$) after ϵ-sampling

1: $\mathcal{D}^\epsilon \leftarrow \emptyset, i \leftarrow 1, j \leftarrow 1$
2: **repeat**
3:     **if** $i \leq N_F$ **then**
4:         $\mathcal{A} \leftarrow \mathcal{F}_i \cup \mathcal{D}^\epsilon$
5:         $i \leftarrow i + 1$
6:     **else**
7:         $\mathcal{A} \leftarrow \mathcal{D}^\epsilon$
8:     **end if**
9:     ϵ-sampling(ϵ, $\mathcal{A}$, $\mathcal{S}$, $\mathcal{D}^\epsilon$)
10:     $\mathcal{F}_j^\epsilon \leftarrow \mathcal{S}$
11:     $j \leftarrow j + 1$
12: **until** $\mathcal{D}^\epsilon = \emptyset$
13: **return**

The new ranking of solutions obtained with ϵ-ranking is reflected in the selection procedure of the algorithm, both during population truncation to obtain the new parent population and during mating to obtain the new offspring population by means of recombination and mutation.

4. Test Problems, Performance Measures and Parameters

In this work we test the performance of the algorithms on multiobjective MNK-Landscapes. A multiobjective MNK-Landscape [2] is defined as a vector function mapping binary strings into real numbers $\boldsymbol{f}(\cdot) = (f_1(\cdot), f_2(\cdot), \cdots, f_M(\cdot)) : \mathcal{B}^N \rightarrow \Re^M$, where M is the number of objectives, $f_i(\cdot)$ is the i-th objective function, $\mathcal{B} = \{0, 1\}$, and N is the bit string length. $\boldsymbol{K} = \{K_1, \cdots, K_M\}$ is a set of integers where K_i ($i = 1, 2, \cdots, M$) is the number of bits in the string that epistatically interact with each bit in the i-th landscape. Each $f_i(\cdot)$ is a non-linear function of $\boldsymbol{x}$ expressed by a Kauffman's NK-Landscape model of epistatic interactions.

We use the hypervolume $\mathcal{H}$ and coverage $\mathcal{C}$ measures [4] to evaluate and compare the performance of the algorithms. The measure $\mathcal{H}$ calculates the volume of the M-dimensional region in objective space enclosed by a set of non-dominated solutions and a dominated reference point. In this work, the reference point is set to $[0.0, \cdots, 0.0]$. Given two sets of non-dominated solutions $\mathcal{A}$ and $\mathcal{B}$, if $\mathcal{H}(\mathcal{A}) > \mathcal{H}(\mathcal{B})$ then set $\mathcal{A}$ can be considered better on convergence and/or diversity of solutions. The coverage $\mathcal{C}$ measure [4] provides complementary information on convergence. Given two sets of non-dominated solutions $\mathcal{A}$ and $\mathcal{B}$, $\mathcal{C}(\mathcal{A}, \mathcal{B})$ gives the fraction of solutions in $\mathcal{B}$ that are dominated at least by one solution in $\mathcal{A}$. Usually $\mathcal{C}(\mathcal{A}, \mathcal{B}) + \mathcal{C}(\mathcal{B}, \mathcal{A}) \neq 1.0$, so both $\mathcal{C}(\mathcal{A}, \mathcal{B})$ and $\mathcal{C}(\mathcal{B}, \mathcal{A})$ are required to understand the degree to which solutions of one set dominate solutions of the other set.

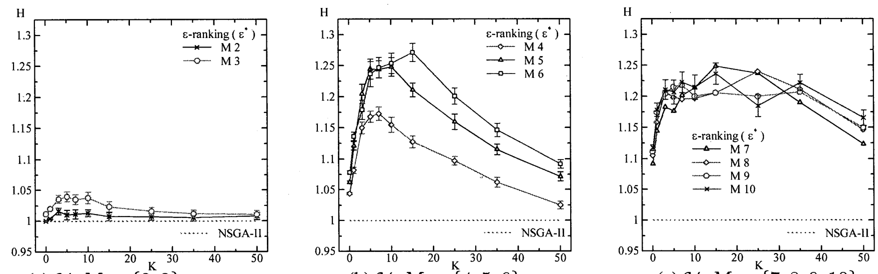
In our experiments we use MNK-Landscapes with $2 \leq M \leq 10$ objectives, $N = 100$ bits, number of epistatic interactions $K = \{0, 1, 3, 5, 10, 15, 25, 35, 50\}$ ($K_1, \cdots, K_M = K$), and *random* epistatic patterns among bits for all objectives. Results presented below show the average performance of the algorithms on 50 different problems randomly generated for each combination of M, N and K. In the plots, error bars show 95% confidence intervals on the mean.

The MOEAs use parent and offspring populations of size 100, two point crossover for recombination with rate $p_c = 0.6$, and bit flipping mutation with rate $p_m = 1/N$ per bit. The number of evaluations is set to 3×10^5. Parameter ϵ for ϵ-ranking is swept in the range $[0.5, 10.0]$ (%) on intervals of 0.5.

5. Experimental Results and Discussion

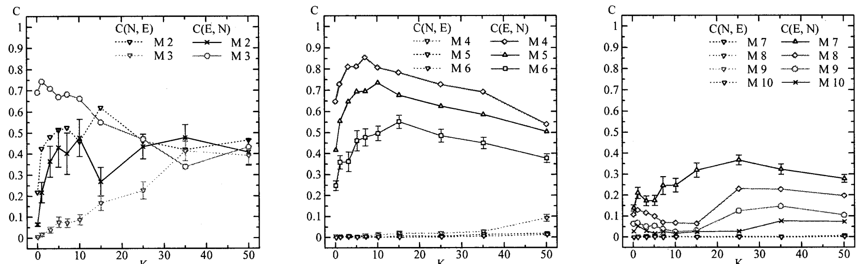
In this section, we discuss the relative gains on performance by ϵ-ranking set with ϵ^* that achieves maximum hypervolume $\mathcal{H}$. **Fig. 1** shows the average ratio $\frac{\mathcal{H}(E)}{\mathcal{H}(N)}$, where E and N denote the set of solutions found by the enhanced NSGA-II with ϵ-ranking (referred as ϵ-ranking for short) and conventional NSGA-II, respectively. As a reference, we include a horizontal line to represent the $\mathcal{H}(N)$ values normalized to 1.0. From this figure, we can see that ϵ-ranking can slightly improve $\mathcal{H}$ on problems with $M = 2$ and $M = 3$ objectives for some values of K (4% improvement or less). On the other hand, for $4 \leq M \leq 10$ objectives, the improvement on $\mathcal{H}$ is remarkable for most values of K (up to 27% improvement). Note that improvements on $\mathcal{H}$ become larger as we increase the number of objectives M from 2 to 6, whereas improvements on $\mathcal{H}$ are similar for $7 \leq M \leq 10$.

Improvements on $\mathcal{H}$ can be due to solutions with better convergence, better diversity, or both. To complement the analyzes of results on $\mathcal{H}$ we also present results using the $\mathcal{C}$ measure. **Fig. 2** shows the average $\mathcal{C}$ values between conventional NSGA-II and ϵ-ranking set with ϵ^*. From this figure, we can be see that $\mathcal{C}(E, N)$ is slightly smaller than $\mathcal{C}(N, E)$ for $M = 2$ and $K \leq 10$, which means that convergence is somewhat worse by ϵ-ranking than conventional NSGA-II. Thus, the slight improvement on $\mathcal{H}$ by ϵ-ranking, ob-



(a) $\mathcal{H}$, $M = \{2, 3\}$ (b) $\mathcal{H}$, $M = \{4, 5, 6\}$ (c) $\mathcal{H}$, $M = \{7, 8, 9, 10\}$

Figure 1: Normalized $\mathcal{H}$, ϵ-ranking set with ϵ^* that achieves maximum $\mathcal{H}(E)$.



(a) $\mathcal{C}$, $M = \{2, 3\}$ (b) $\mathcal{C}$, $M = \{4, 5, 6\}$ (c) $\mathcal{C}$, $M = \{7, 8, 9, 10\}$

Figure 2: $\mathcal{C}$ between NSGA-II and ϵ-ranking set with ϵ^* that achieves maximum $\mathcal{H}(E)$.

served in the same region $M = 2$ and $0 \leq K \leq 10$ as shown in **Fig. 1**, is at the expense of a slight deterioration on convergence. In the case of $M = 3$ objectives, we can see that $\mathcal{C}(E, N)$ is considerably greater than $\mathcal{C}(N, E)$ for $0 \leq K \leq 25$, which means that convergence is better by ϵ-ranking than conventional NSGA-II. Thus, we can conclude that a better convergence of solutions contributes to the improvement on $\mathcal{H}$ by ϵ-ranking, as shown in **Fig. 1** for $0 \leq K \leq 25$.

For $4 \leq M \leq 10$ a clear trend can be seen. $\mathcal{C}(N, E)$ is close to 0.0 for most K and M. This indicates that there are almost no solutions by conventional NSGA-II that dominate solutions by ϵ-ranking. On the other hand, the values of $\mathcal{C}(E, N)$ are very high for 4 objectives (in the range 0.55-0.85) and reduce gradually as we increase M up to 10 objectives (in the range 0.01-0.08). This suggests that a better convergence of solutions contributes to the increases of $\mathcal{H}$ by ϵ-ranking on $M = 4$ problems. As we increase M, gains on diversity gradually become more significant than gains on convergence as the reason for the remarkable improvement of $\mathcal{H}$ on $5 \leq M \leq 10$.

6. Conclusions

In this work, we have proposed ϵ-ranking to enhance selection of MOEAs aiming to improve their performance on *many* objectives problems. The proposed method uses a randomized sampling procedure to increase selection probabilities of some of the solutions, while trying to keep a uniform search effort towards the different zones of objective space represented in the instantaneous population. We enhanced NSGA-II with the proposed method showing that convergence and diversity of the obtained solutions can improve remarkably on MNK-Landscapes with $3 \leq M \leq 10$ objectives and $0 \leq K \leq 50$ epistatic interactions per bit. As future works, we would like pursue adaptive methods to control the parameter ϵ.

Acknowledgment

This study was performed through Special Coordination Funds for Promoting Science and Technology of the Ministry of Education, Culture, Sports, Science and Technology, of the Japanese Government.

References

[1] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*, John Wiley & Sons, Chichester, West Sussex, England, 2001.

[2] H. Aguirre and K. Tanaka, "Working Principles, Behavior, and Performance of MOEAs on MNK-Landscapes", *European Journal of Operational Research*, Elsevier, vol. 181(3), pp. 1670-1690, September 2007.

[3] M. Laumanns, L. Thiele, K. Deb and E. Zitzler, "Combining Convergence and Diversity in Evolutionary Multi-objective Optimization", *Evolutionary Computation*, Vol.10, No.3, pp.263-282, Fall 2002.

[4] E. Zitzler, *Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications*, PhD thesis, Swiss Federal Institute of Technology, Zurich, 1999.