A survey of nearest-better clustering in swarm and evolutionary computation

Luo, W.; Lin, X.; Zhang, J.; Preuss, M.

A Survey of Nearest-Better Clustering in Swarm and Evolutionary Computation

1st Wenjian Luo
School of Computer Science and Technology
Harbin Institute of Technology
Shenzhen 518000, Guangdong, China
luowenjian@hit.edu.cn

2nd Xin Lin
School of Computer Science and Technology
University of Science and Technology of China
Hefei 230027, Anhui, China
iskcal@mail.ustc.edu.cn

3rd Jiajia Zhang
School of Computer Science and Technology
Harbin Institute of Technology
Shenzhen 518000, Guangdong, China
zhangjiajia@hit.edu.cn

4th Mike Preuss
Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
The Netherlands
m.preuss@liacs.leidenuniv.nl

2021 IEEE Congress on Evolutionary Computation (CEC) | 978-1-7281-8393-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/CEC45853.2021.9505008

***Abstract*—Nearest-Better Clustering (NBC) is an emergent niching technique in Swarm and Evolutionary Computation for optimization, which does not need to fix the number or radius of clusters in advance. The key idea of NBC is to first link each individual to its nearest better neighbor to form a spanning tree of all individuals in the population, and then partition all individuals into clusters by deleting the longer edges in the spanning tree. In this paper, a survey on the Nearest-Better Clustering algorithms and applications in multimodal and dynamic optimization is provided. First, the basic NBC algorithm is introduced. Second, the improvements of the basic NBC are detailed. Third, multimodal and dynamic optimization algorithms powered by NBC are enlisted and discussed.**

***Index Terms*—Evolutionary Algorithms, Swarm Intelligence, Nearest Better Clustering**

I. Introduction

In the real world, lots of problems are optimization problems where an optimization algorithm is expected to find the globally optimal solution and return it to the decision maker. Among these algorithms, the population-based stochastic search algorithms in the swarm and evolutionary computation field, such as differential evolution [1], [2], [3], [4], particle swarm optimization [5], [6], [7] and genetic algorithms [8], [9], [10], have been widely adopted to solve various types of optimization problems.

In the field of multimodal optimization, a multimodal optimization problem (MMOP) has multiple optima (often with several acceptable solutions) which should be found by the algorithm, and the decision maker selects the preferred solution among them [11]. Here, acceptable solution means that the quality of the solutions is near to what can be achieved given the available resources. Besides, in the field of dynamic optimization [12], [13], because the fitness landscape of a dynamic optimization problem (DOP) may be changed when an environmental change occurs, a local optima peak in the previous environment may transfer to a global peak after the environmental change. Therefore, many dynamic optimization algorithms adopt niching methods to catch all the global and promising local optima in the current environment in order to increase the ability to rapidly locate the area of the (or one) global optimum after the change [12], [14], [15].

However, evolutionary algorithms as well as swarm intelligence methods have initially been designed for finding one single best solution. As the evolution progresses, populations/swarms eventually converge to a single search space area, so that detecting multiple optimal solutions at the same time is not possible any more. Therefore, niching methods have been proposed to enhance the search ability of finding multiple optima simultaneously [11], [16], [17].

For example, in [18], Li proposed species-based particle swarm optimization (SPSO) to solve multimodal optimization problems, and then in [19] he proposed species-based differential evolution (SDE). Both algorithms are based on the speciation method to divide the population into multiple species. The speciation method adopts a species radius to cluster these individuals in the population. For each individual, from the best to the worst, if the distances of each individual to all the best ones of existing species are larger than the species radius, the current solution becomes the best of a new species. Otherwise, it is grouped into the nearest species. For another example, Qu and his colleagues [20] proposed neighborhood mutation based differential evolution, where each species is formed by m nearest individuals, and m is a preset parameter.

Nearest-better clustering (NBC) [21] is a very promising niching technique in the fields of multimodal optimization and dynamic optimization. NBC does not require the prior knowledge of the optimization problem and has only one

This study is supported by by the National Key Research and Development Program of China (No. 2020YFB2104003) and the National Natural Science Foundation of China (No. 61573327). *(Corresponding author: Wenjian Luo.)*

parameter to be set. NBC assumes that the distance between two optimal solutions in different peaks are much larger than the weighted average distance from each one to its nearest-better neighbor. Based on this assumption, NBC is designed and has two primary steps as follows [21]. First, add an edge between each individual and its nearest-better neighbor, and then a spanning tree is constructed. Second, by cutting the longer edges, the spanning tree are divided into several components, and each component represents one species. In this paper, we make a full description about the basic process and the improvements of the nearest-better clustering (NBC) method.

The rest of this paper is organized as follows. Section II gives a detail description about the basic NBC. Section III discusses the shortcomings of the basic NBC and the corresponding improvements. The applications of NBC in different evolutionary optimization fields are illustrated in Section IV. Section V compares NBC with some related niching methods and makes a discussion. Finally, a brief conclusion is depicted in Section VI.

II. Basic Algorithm of Nearest-Better Clustering

Before the basic algorithm of NBC is given, the nearest-better neighbor is explained as follows. For individual x, its nearest-better neighbor is defined as the solution which is the nearest to x and whose fitness values are better than x. That is,

$$NBI(x) = \underset{x_i \in P,\ f(x_i) > f(x)}{\arg\min} d(x_i, x) \tag{1}$$

Where $NBI(x)$ is the nearest-better individual of individual x, P is the population, $d(x_i, x)$ is the Euclidean distance between x_i and x, and $f(.)$ is the fitness function, and suppose that the individual with a greater fitness value is better.

Considering a maximization problem, the basic NBC is shown in Algorithm 1 [21], [22]. First, all the individuals in the population are sorted by fitness values from the best to the worst. Second, for each individual except the best solution (the best one has no nearest-better neighbor), its nearest-better neighbor is found, and an edge between them is added, and the length of the edge is the Euclidean distance. Consequently, all individuals are connected and a spanning tree T is constructed. Third, the weighted mean distance $mean\_dis$ of all the edges in T is calculated, where $mean()$ is the average function and φ is a predefined parameter usually set to 2.0. Fourth, each edge in T is traversed and checked whether its length is greater than $mean\_dis$. If the length of an edge is greater than $mean\_dis$, the edge will be cut off. Finally, after the longer edges are cut off, the spanning tree T becomes a forest in which each sub-tree represents a sub-population (called species or cluster in this paper), and the best individual of each sub-tree is the seed of the corresponding species.

The nearest-better clustering method (NBC) is a powerful niching technique in evolutionary multimodal optimization and evolutionary dynamic optimization. However, there are some disadvantages in the basic NBC, and some improvements have been done, which will be introduced in the following section.

Algorithm 1 Nearest-Better Clustering

Input: P (population), F (fitness values), φ (the weight factor)
Output: $Species$ (the partitioning scheme)

Sort all individuals in the population P by fitness values F in descending order;
Create an empty tree T and an empty distance set $dists$;
for each individual x_i **in** P **do**
  if x_i is the best one **then**
    Continue;
  end if
  Find the nearest-better neighbor $NBI(x_i)$;
  Connect the edge between x_i and $NBI(x_i)$ in T;
  Save the distance $d(x_i, NBI(x_i))$ in $dists$;
end for
$mean\_dis = \varphi * mean(dists)$;
for each edge e **in** T **do**
  if $e.dis > mean\_dis$ **then**
    Cut off e;
  end if
end for
for each tree t in Forest T **do**
  Set the best individual in t as the seed;
  Regard the individuals in t as a species and store them in $Species$;
end for

III. Improvements

The basic NBC is a simple and effective niching technique and it has two significant advantages when the optimization algorithm divides the population into multiple species. The first one is that the basic NBC does not need of any prior knowledge of the optimization problem, such as the number of global optima in the fitness landscape. The second advantage is that there is only one parameter to be set. The parameter φ controls the cutting threshold, only the edges whose length are greater than the weighted average length should be cut. In most work, φ is set to 2.0 as a default value. It is possible to make the methods a bit more sensitive by reducing φ (note that a $\varphi = 1$ would effectively mean that on average every second connection would be cut, which is certainly too much for a reasonable clustering), or less sensitive by increasing φ. There are some guidelines in [23] on how to set the parameter, depending on the dimensionality of the problem, but these are purely based on experiments, there is no theory to build on.

As with the heuristic setting of φ, NBC also has its weeknesses, and some improvements have been suggested in order to cure these. We describe these in the following.

A. How to set parameter φ

Parameter φ is the only parameter in the basic NBC. Therefore, it is very important to study its impact on performance. In general, parameter φ is set to a constant 2.0. However, this value is an empirical parameter, and it still needs to be investigated. Thus, the question on how to set this parameter leads to some possible modifications of the method.

In [24], Luo *et al.* performed an in-depth study on NBC, and proposed strategy f to set the value of φ. Different from the fixed φ value in the basic NBC, this work considers that there is a large fluctuation in the performances of the basic NBC in the face of different problems or at different evolutionary stages. To improve its performance and try not to introduce too many parameters related to the specific problems, they set φ to a random value between lb and ub, which is shown in Formula (2). The lower bound lb and upper bound ub are set to 1.25 and 2.25, respectively.

$$\varphi = rand(lb, ub) \tag{2}$$

In [25], Lin *et al.* proposed a method named NBC-minsize, where the parameter φ is set to a very small value of 1.0 in order to separate individuals in the areas with multiple close optimal individuals. However, 1.0 is too small, by which NBC generates a huge amount of species, and many species only have a very small number of individuals. Therefore, to ensure the minimum size of species, a new cutting strategy is added in [25], which will be discussed in the following section III-B.

B. Which edges should be cut

Due to inaccuracy of the partition results only depending on parameter φ, some studies have paid attention on adding some cutting rules.

In [26], Preuss improved the NBC to achieve better performance on 5-20 dimensional optimization problems. Besides the cutting rules by the weighted average distance which depends on parameter φ, another cutting rule based on the incoming edges is added. Because each edge connects an individual with its nearest-better neighbor, an individual (except the best one) must have only one edge to its nearest-better neighbor (called an outgoing edge), and an individual in the population may have multiple edges which regard itself as others' nearest-better neighbor (called incoming edges). Preuss considered an individual has an outgoing edge and at least 3 incoming edges might be a seed of a new cluster, and checked the ratio between the length of the outgoing edge and the average length of incoming edges. If the ratio is greater than a parameter b, the outgoing edge will be cut. It should be noted that the edges with length greater than the weighted average distance will also be cut off in [26].

In [25], the proposed NBC-minsize adds an additional rule, which set the minimum size of species, to avoid generating very small species. First, all the edges are sorted by their lengths in descending order. Then, for each edge, if its length is greater than the weighted average distance and the sizes of two clusters generated by cutting off this edge are both greater than a threshold $minsize$, this edge should be cut off. In [25], during evolution, the parameter $minsize$ is linearly increased with evolutionary iterations until it reaches a maximum number. With a small value 1.0 of parameter φ, it can be seen that more clusters are generated in the early stage while less ones in the latter stage, which has a powerful balance between exploration and exploitation.

C. Which individuals should be used to construct the spanning tree

In the basic NBC, all individuals participate in the process of constructing a spanning tree. However, in [24], Luo *et al.* believed that constructing a spanning tree by all individuals is not the best way for population partitioning. During evolution, there often exist some individuals which have much worse fitness. These outliers have a bad influence on the performance of NBC. They are scattered in the whole decision space, and usually gathered into multiple clusters with a small number of individuals. The size of these clusters is too small to evolve so that it wastes the computational resources. Meanwhile, these poor individuals often have a large distance to their nearest-better neighbors, which results in a significant increase on the weighted average distance, while a larger threshold cannot properly separate the individuals in multiple close peak areas. Therefore, the authors propose strategy p to reduce the number of individuals used in NBC. That is, only a part of individuals with good fitness values are used to construct the spanning tree in NBC. This strategy is also adopted in [27].

In detail, two implementations of strategy p are proposed in [24], which are described as follows.

1) A fixed ratio is adopted. That is, the best $(1-r)*N$ individuals are used to construct the spanning tree of NBC, where N is the population size and r is is a parameter. In [24], r is set to a fixed value 0.1.

2) The number of individuals used to construct the spanning tree of NBC is adaptive. At the first generation of the evolutionary process, the best $(1-r)*N$ individuals are used. Then, the number of the re-initialized individuals in each generation is recorded. Suppose the number of individuals re-initialized in the previous generation is $numInitial$. If $numInitial$ is greater than 0, the worst $(numInitial-1)$ individuals are not adopted to construct the spanning tree. Otherwise, all the individuals are selected. This is because the main algorithm in [24] is designed for dynamic optimization, and the random immigrant strategy is often used to enhance the exploration ability of the evolutionary dynamic optimization algorithms.

Besides, in [28], Li *et al.* introduced a weighted gradient and distance-based clustering method (WGraD) to overcome the shortcoming of NBC. In WGraD, when constructing the spanning tree, each individual is connected to the other one based on the weighted gradient and weighted distance information.

D. Seeds-based partitioning

In the basic NBC, the spanning tree is constructed and used to form the species. However, this might not be the best way in some cases. In [24], Luo *et al.* described a situation which is called "*long tail*" phenomenon. Different global peaks have different width, so there are more individuals on wide peaks, and fewer individuals on narrow peaks. The "*long tail*" phenomenon is that some individuals which belong to the wide peaks might be nearer to the optima of a narrow peak instead of the wide one. Due to a too large number of individuals in the species locating at the wide peak, these "long tail" individuals have less contribution to search the wide peak. It is better to put these individuals in the species searching the global optima of the narrow peak.

Based on the above observation, Luo *et al.* [24] designed strategy s to improve the NBC. The implementation is very simple and described as follows. First, the basic NBC is adopted only to obtain the seeds rather than dividing the population. Then, each individual is traversed and put into the cluster with the seed nearest to it.

In [29], in order to reduce the time cost of NBC, Bu *et al.* proposed a two-stage clustering method named κ*-NBC+k-NSC*. First, κ*-NBC* is used to divide the population, where κ is set 2. If the number of species keeps unchanged in k generation, k-NSC is adopted to divide the population, where k is set 3. For κ*-NBC*, an edge in the spanning tree, from x to its nearest-better neighbor $NBI(x)$, will be deleted if the following two conditions are both satisfied: (1) its length is larger than the weighted distance; (2) x is the nearest-better neighbors of at least κ individuals. For k-NSC (the k-step nearest seeds clustering), the seeds in the previous generation are used as the initial seeds in the current generation, and each individual is grouped into the cluster with the seed nearest to it. Then, the best individual in each species is selected as the new seed, and every individual is reassigned into the cluster with the nearest seed. The final species are obtained by repeating the above step k-1 times .

E. Other improvements

In [25], NBC is executed twice in one evolutionary generation. First, the population is divided by NBC-minsize. Second, in each species obtained by NBC-minsize, the basic NBC is executed once again. The second execution of NBC is only used to find the seeds in a species which is called *keypoints*, and does not divide the species again. These *keypoints* are all regarded as the most promising individuals in the species. Furthermore, in order to utilize the found *keypoints*, the DE/best/1 and DE/best/2 mutation operators in the DE framework are replaced by DE/keypoint/1 and DE/keypoint/2, respectively, which are shown as follows.

$$\begin{aligned} \boldsymbol{z}_i &= \boldsymbol{x}_{kp} + F * (\boldsymbol{x}_{r1} - \boldsymbol{x}_{r2}), \\ \boldsymbol{z}_i &= \boldsymbol{x}_{kp} + F_1 * (\boldsymbol{x}_{r1} - \boldsymbol{x}_{r2}) + F_2 * (\boldsymbol{x}_{r3} - \boldsymbol{x}_{r4}), \end{aligned} \tag{3}$$

where $\boldsymbol{x}_{kp}$ is one randomly selected *keypoint*, $\boldsymbol{x}_{r1}$, $\boldsymbol{x}_{r2}$, $\boldsymbol{x}_{r3}$ and $\boldsymbol{x}_{r4}$ are randomly selected individuals in the species, $\boldsymbol{z}_i$ is the i-th offspring individual.

IV. APPLICATIONS

The nearest-better clustering method performs well in the field of evolutionary multimodal optimization. Besides, in evolutionary dynamic optimization, because finding all the global and promising local optima is very beneficial to rapidly respond to environmental changes, NBC has also been used to in some evolutionary dynamic optimization algorithms, recently. In this section, the related work about NBC in evolutionary multimodal optimization and evolutionary dynamic optimization are discussed.

A. Evolutionary multimodal optimization

NBC is initially designed for evolutionary multimodal optimization, and it is easily embedded into most of the swarm and evolutionary algorithms.

In [25], Lin *et al.* proposed a multimodal optimization algorithm called FBK-DE focusing on the species formulation, species balance and keypoints in species. In each generation, after the population partition by NBC-minsize and resizing the species, each species is regarded as a sub-population and evolved by differential evolution. FDK-DE adopts the basic mutation operators (i.e., DE/rand/1 and DE/rand/2) as well as improved operators (i.e., DE/keypoint/1 and DE/keypoint/2) to generate the offspring. In [30], in order to effectively solve multimodal optimization problems, Dowlatshahi *et al.* embedded the topological nearest-better and distance-based nearest-better strategy into the gravitational search algorithm (GSA).

For the evolution strategy with covariance matrix adaptation (CMA-ES), because the covariance matrix and the step size are metadata related to the population, the way to combine NBC and CMA-ES might be a bit different. Preuss [22] proposed a novel algorithm embedding NBC into CMA-ES. First, the algorithm randomly generates a population of individuals and evaluates them. Then, NBC is executed to divide the population into multiple sub-populations. Each sub-population represents a population in CMA-ES and is evolved in one generation by CMA-ES. After generating offspring of all sub-populations, the offspring are separated by NBC into next sub-populations. The metadata of each next sub-population is derived from the parent sub-population which has the most parent individuals.

Combing NBC with other clustering methods could be an effective way to improve the performance of evolutionary algorithms for solving MMOPs. In [31], Pereira *et al.* adopted the DM (Detect Multimodal, originated from the hill-valley method [32], [33]) method to guide the population partition after the NBC generates the spanning tree. Another combination of NBC and DM is used in [34], where the DM method is adopted to filter the clusters generated by NBC. Similarly, in [35], the proposed nearest-better separation evolutionary algorithm (NBSEA) adopted DM to check if two seeds generated by NBC locates at the same peak. If they are in the same peak, two clusters with corresponding seeds are unified.

In [36], Preuss *et al.* adopted the proposed algorithm in [22], [26] to search for the multiple optimal positions of virtual cameras in 3D applications, especially in the 3D games. In [37], Preuss *et al.* studied the game content generation problems by NEA2, which is based on NBC. In [38], a two-stage evolutionary algorithm is designed to solve Nash Equilibrium problems, where both NBC and DM are adopted to cluster the individuals during the searching procedure.

In addition, in [27], Luo *et al.* adopted a clonal selection algorithm as the basic optimizer and the improved version of NBC in order to solve many-modal optimization problems. Many-modal optimization problems have properties similar to the MMOPs, but the number of the global optima (sometimes including acceptable local optima) is much larger.

B. Evolutionary dynamic optimization

Recently, some dynamic optimization algorithms utilize NBC to find all the global and local optima in an environment, in order to quickly track the global optima after an environmental change occurs [39].

Luo *et al.* [24] adopted NBC to identify multiple species in different global/local optimal areas, which is beneficial for PSO (particle swarm optimization) to track the global optima under dynamic environments. When the environment changes, the memory preserves the best 5 seeds generated by NBC of current environment, and select 3 individuals in the memory to replace the worst three individuals in the population.

Bu *et al.* [29] proposed a species-based memory-enhanced differential evolution (SMEDE) to solve the optimal power flow (OPF) problems under double-sided uncertainties, where an improved NBC named *κ-NBC+k-NSC* was proposed.

In [40], a dynamic multimodal optimization algorithm based on the clonal selection algorithm was proposed, where NBC was adopted to divide the population to search for the multiple promising areas in each environment. The dynamic multimodal optimization problems (DMMOPs) have the dynamic and multimodal properties. Each environment of DMMOPs has multiple global optimal solutions which the algorithms need to find. Later, an empirical study was done in [41] to compare different dynamic response strategies and basic optimizers in solving dynamic constrained multimodal optimization problems, where the basic NBC is adopted to identify the species in the evolutionary population.

V. Discussion

A. Comparison with other clustering methods

Some work has been done to discuss the performance comparison of NBC and other clustering methods including the topographical selection (TS) method [42], the Jarvis-Patrick (JP) clustering method [43], and the nearest-better neighbor clustering (NBNC) method [44].

Different from NBC, TS [42] directly construct the forest instead of the spanning tree to obtain the clusters. Each individual connects to its k nearest neighbors and the direction of the edge is from the worse individual to the better solution. All the individuals having no outdegree are considered to be the seeds. In [45], Wessing *et al.* compared the performance between TS and NBC. According to the experimental results, the authors conclude that both methods have their own advantges and individual parameter settings, and obtaining some information about the specific problems is helpful for the parameter settings. The authors also suggested that it may be a good idea to combine NBC and TS.

JP [43] is based on the distance of several nearest neighbors instead of the nearest-better neighbor. In JP, each individual finds its J nearest neighbors and store them in a set. Two individuals considered to be in the same cluster should satisfy two conditions: (1) both solutions should be in the nearest neighbors set of each other; (2) there are at least K individuals in the intersection of two nearest neighbor set. Two parameters J and K should be predefined and have a great impact on the performance. In [46], JP and NBC are compared, and experimental results shows that NBC has a slight advantage against to JP. In [47], Stoean *et al.* compared JP, NBC and topological multimodality estimator (TME) based on the detect-multimodal method, and found that NBC has a powerful performance on low fitness evaluation calls, while TME is better on high fitness evaluation calls.

NBNC [44] utilizes the information of distances about the nearest neighbor and the nearest-better neighbor. To achieve a reasonable partition scheme, NBNC first obtains the distances from individuals to their corresponding nearest neighbor. The next step is to calculate the distance threshold (weighted average nearest distance). Then, each individual from the best to the worst looks for the nearest-better neighbor within the distance threshold, and if the nearest-better neighbor exists, these two individuals are connected. Finally, the whole population will be transformed into a forest with multiple trees, and each tree represents a cluster. In [44], Luo *et al.* compares the performance of NBNC and NBC, by embedding them into a hybrid algorithm of an evolutionary strategy and particle swarm optimization, respectively. Experimental results show that compared with NBC, NBNC has better performance on the multimodal optimization problems.

B. Nearest-better tree

Recently, some work introduced a definition of the nearest-better tree, which is constructed by connecting each individual to its nearest-better neighbor, and whose root is the best individual in the population.

In [48], Maree *et al.* proposed a hierarchical Gaussian mixture learning (HGML) algorithm to cluster the individuals in the population. After dividing the population into multiple clusters by the proposed method based on Gaussian mixture model (GMM), the authors consider that it is necessary to merge some clusters. However, it may be not accurate to merge clusters only in terms of distance. Therefore, the algorithm uses the nearest-better tree for this purpose. That is, each cluster is regarded as a node and a cluster is connected to its nearest-better one, and then a spanning tree is formed. In order to reduce the computational complexity, the algorithm selects two nearest neighbor clusters in the constructed tree for merging.

Another work by Chen *et al.* [49] adopted two-leveled Gaussian mixture model to solve the problems with multimodal and rugged environments. In their work, the nearest-better tree is used to obtain the possible local optimal solution, which is constructed in the same way as the basic NBC. However, when partitioning the population according to the nearest-better tree, Chen *et al.* considered that if one node has a child closer to the root of tree than itself, it is more likely to be the root of a new sub-tree.

C. Nearest-better neighbor information

In NBC, the nearest-better neighbor information is adopted to construct the spanning tree. Nearest-better neighbor information (i.e., the distance) has also been used in multiple fields, such as the exploratory landscape analysis (ELA).

Exploratory landscape analysis [50], [51] is a collection of techniques to acquire the information about the properties of unknown optimization problems. Obtaining some knowledge about a problem may be helpful for solving the problem. ELA has many features to describe the fitness landscape of a problem, such as the global structure, multimodality, separability, plateaus, and so on. Among them, the information about the nearest-better neighbor is often adopted to analyze the multimodality of the landscape. In [52], Kerschke *et al.* proposed 5 indicators to describe the intensity of multimodality with two types of distances, i.e., the distance to the nearest-better neighbor and the distance to the nearest neighbor. For example, the first indicator is equal to the ratio of the standard deviations of two distance sets. It is believed that this value is very large on a landscape with the multimodal property and almost equal to 1 on an unimodal problem. That is because the distances to nearest-better neighbors may have a dramatic change on the multimodal optimization problems, which makes the corresponding standard deviation much larger. In addition, in [53], Kerschke *et al.* introduced a R-package named FLACOO for feature-based landscape analysis of continuous and constrained optimization problems, and the nearest-better clustering features are included in FLACOO.

Some works convert the original single optimization problems to multiobjective optimization problems, while nearest-better neighbor information is regarded ass one objective. For example, in [54], Bischl *et al.* adopted several types of distance as the second objective and proposed a multiobjective optimization algorithm to solve parallel model-based optimization problems. The experimental results show that the distance to the nearest-better neighbor has a more powerful performance on average. Wessing *et al.* [55] tried to solve MMOPs by the multiobjective optimization algorithms. To balance exploitation and exploration, two conflicting objectives are designed, i.e., the fitness value and the average distance of at most k nearest-better neighbor. Experimental results showed that the algorithm has competitive performance. In [56], Wessing *et al.* compared the performance of the nearest neighbor and the nearest-better neighbor by multiobjective optimization algorithms for solving MMOPs, and illustrated that the nearest-better neighbor strategy wins.

VI. Conclusions

Nearest-better clustering is a simple but effective niching technique. In this paper, firstly, the detailed process of nearest-better clustering is illustrated. Then, some disadvantages and corresponding improvements about the basic nearest-better clustering are given. Next, the applications of nearest-better clustering in the multimodal optimization and dynamic optimization are discussed in detail. Finally, the work about comparing NBC with other clustering methods, and the work about the nearest-better tree, as well as the nearest-better neighbor information used in exploratory landscape analysis and evolutionary multiobjective optimization algorithms, are introduced. Overall, nearest-better clustering and its variants have demonstrated the competitive performance on locating multiple optimal areas and identifying the species in the population.

References

[1] R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[2] S. Das and P. N. Suganthan, "Differential evolution: A survey of the state-of-the-art," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 1, pp. 4–31, 2010.

[3] F. Neri and V. Tirronen, "Recent advances in differential evolution: a survey and experimental analysis," *Artificial Intelligence Review*, vol. 33, no. 1-2, pp. 61–106, 2010.

[4] S. Das, S. S. Mullick, and P. N. Suganthan, "Recent advances in differential evolution–an updated survey," *Swarm and Evolutionary Computation*, vol. 27, pp. 1–30, 2016.

[5] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*. Ieee, 1995, pp. 39–43.

[6] Y. Zhang, S. Wang, and G. Ji, "A comprehensive survey on particle swarm optimization algorithm and its applications," *Mathematical Problems in Engineering*, vol. 2015, 2015.

[7] K. Kameyama, "Particle swarm optimization-a survey," *IEICE Transactions on Information and Systems*, vol. 92, no. 7, pp. 1354–1361, 2009.

[8] J. R. Sampson, "Adaptation in natural and artificial systems (john h. holland)," 1976.

[9] L. Davis, "Handbook of genetic algorithms," 1991.

[10] T. Li, G. Shao, W. Zuo, and S. Huang, "Genetic algorithm for building optimization: State-of-the-art survey," in *Proceedings of the 9th International Conference on Machine Learning and Computing*, 2017, pp. 205–210.

[11] S. Das, S. Maity, B.-Y. Qu, and P. N. Suganthan, "Real-parameter evolutionary multimodal optimization—a survey of the state-of-the-art," *Swarm and Evolutionary Computation*, vol. 1, no. 2, pp. 71–88, 2011.

[12] M. Mavrovouniotis, C. Li, and S. Yang, "A survey of swarm intelligence for dynamic optimization: Algorithms and applications," *Swarm and Evolutionary Computation*, vol. 33, pp. 1–17, 2017.

[13] S. Yang, "Evolutionary computation for dynamic optimization problems," in *Proceedings of the Companion Publication of the 2015 Annual Conference on Genetic and Evolutionary Computation*, 2015, pp. 629–649.

[14] T. T. Nguyen, S. Yang, and J. Branke, "Evolutionary dynamic optimization: A survey of the state of the art," *Swarm and Evolutionary Computation*, vol. 6, pp. 1–24, 2012.

[15] C. Cruz, J. R. González, and D. A. Pelta, "Optimization in dynamic environments: a survey on problems, methods and measures," *Soft Computing*, vol. 15, no. 7, pp. 1427–1448, 2011.

[16] X. Li, M. G. Epitropakis, K. Deb, and A. Engelbrecht, "Seeking multiple solutions: an updated survey on niching methods and their applications," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 4, pp. 518–538, 2016.

[17] Y. Mehmood, N. Aziz, F. Riaz, H. Iqbal, and W. Shahzad, "Pso-based clustering techniques to solve multimodal optimization problems: A survey," in *2018 1st International Conference on Power, Energy and Smart Grid (ICPESG)*. IEEE, 2018, pp. 1–6.

[18] X. Li, "Adaptively choosing neighbourhood bests using species in a particle swarm optimizer for multimodal function optimization," in *Proceedings of Genetic and Evolutionary Computation Conference*. Springer, 2004, pp. 105–116.

[19] ——, "Efficient differential evolution using speciation for multimodal function optimization," in *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation*, 2005, pp. 873–880.

[20] B.-Y. Qu, P. N. Suganthan, and J.-J. Liang, "Differential evolution with neighborhood mutation for multimodal optimization," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 5, pp. 601–614, 2012.

[21] M. Preuss, L. Schönemann, and M. Emmerich, "Counteracting genetic drift and disruptive recombination in (μ^+,λ)-EA on multimodal fitness landscapes," in *Proceedings of the 7th annual conference on Genetic and evolutionary computation*, 2005, pp. 865–872.

[22] M. Preuss, "Niching the CMA-ES via nearest-better clustering," in *Proceedings of the 12th annual conference companion on Genetic and evolutionary computation*, 2010, pp. 1711–1718.

[23] ——, *Multimodal Optimization by Means of Evolutionary Algorithms*, ser. Natural Computing Series. Springer, 2015. [Online]. Available: https://doi.org/10.1007/978-3-319-07407-8

[24] W. Luo, J. Sun, C. Bu, and R. Yi, "Identifying species for particle swarm optimization under dynamic environments," in *Proceedings of 2018 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2018, pp. 1921–1928.

[25] X. Lin, W. Luo, and P. Xu, "Differential evolution for multimodal optimization with species by nearest-better clustering," *IEEE Transactions on Cybernetics*, 2019.

[26] M. Preuss, "Improved topological niching for real-valued global optimization," in *Proceedings of European Conference on the Applications of Evolutionary Computation*. Springer, 2012, pp. 386–395.

[27] W. Luo, Y. Qiao, X. Lin, P. Xu, and M. Preuss, "Many-modal optimization by difficulty-based cooperative co-evolution," in *Proceedings of 2019 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2019, pp. 1907–1914.

[28] Y. Li, J. Yu, and H. Takagi, "Niche method complementing the nearest-better clustering," in *Proceedings of 2019 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2019, pp. 3065–3071.

[29] C. Bu, W. Luo, T. Zhu, R. Yi, and B. Yang, "Species and memory enhanced differential evolution for optimal power flow under double-sided uncertainties," *IEEE Transactions on Sustainable Computing*, 2019.

[30] M. B. Dowlatshahi, V. Derhami, and H. Nezamabadi-pour, "Gravitational search algorithm with nearest-better neighborhood for multimodal optimization problems," 2019.

[31] M. W. Pereira, G. S. Neto, and M. Roisenberg, "A topological niching covariance matrix adaptation for multimodal optimization," in *Proceedings of 2014 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2014, pp. 2562–2569.

[32] R. K. Ursem, "Multinational evolutionary algorithms," in *Proceedings of the 1999 Congress on Evolutionary Computation (CEC)*, vol. 3. IEEE, 1999, pp. 1633–1640.

[33] S. Maree, T. Alderliesten, D. Thierens, and P. A. Bosman, "Real-valued evolutionary multi-modal optimization driven by hill-valley clustering," in *Proceedings of the Genetic and Evolutionary Computation Conference*, 2018, pp. 857–864.

[34] M. Preuss, C. Stoean, and R. Stoean, "Niching foundations: basin identification on fixed-property generated landscapes," in *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, 2011, pp. 837–844.

[35] C. Stoean, M. Preuss, T. Bartz-Beielstein, and R. Stoean, *A new clustering-based evolutionary algorithm for real-valued multimodal optimization*. TU, Algorithm Engineering, 2009.

[36] M. Preuss, P. Burelli, and G. N. Yannakakis, "Diversified virtual camera composition," in *Proceedings of European Conference on the Applications of Evolutionary Computation*. Springer, 2012, pp. 265–274.

[37] M. Preuss, A. Liapis, and J. Togelius, "Searching for good and diverse game levels," in *Proceedings of 2014 IEEE Conference on Computational Intelligence and Games*. IEEE, 2014, pp. 1–8.

[38] H. Li, Z. Wu, and G. Zhaolu, "Two-stage multimodal optimization algorithm for solving nash equilibrium," *Journal of Wuhan University(Natural Science Edition)*, vol. 62, no. 5, pp. 444–450, 2016.

[39] W. Luo, B. Yang, C. Bu, and X. Lin, "A hybrid particle swarm optimization for high-dimensional dynamic optimization," in *Proceedings of Asia-Pacific Conference on Simulated Evolution and Learning*. Springer, 2017, pp. 981–993.

[40] W. Luo, X. Lin, T. Zhu, and P. Xu, "A clonal selection algorithm for dynamic multimodal function optimization," *Swarm and Evolutionary Computation*, vol. 50, p. 100459, 2019.

[41] X. Lin, W. Luo, Y. Qiao, P. Xu, and T. Zhu, "Empirical study of population-based dynamic constrained multimodal optimization algorithms," in *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2019, pp. 722–730.

[42] A. Törn and S. Viitanen, "Topographical global optimization," *Recent advances in Global Optimization*, pp. 384–398, 1992.

[43] R. A. Jarvis and E. A. Patrick, "Clustering using a similarity measure based on shared near neighbors," *IEEE Transactions on Computers*, vol. 100, no. 11, pp. 1025–1034, 1973.

[44] W. Luo, Q. Yingying, X. Lin, and M. Preuss, "Hybridizing niching, particle swarm optimization and evolution strategy for multimodal optimization," *IEEE Transactions of Cybernetics*, 2020.

[45] S. Wessing, G. Rudolph, and M. Preuss, "Assessing basin identification methods for locating multiple optima," in *Advances in Stochastic and Deterministic Global Optimization*. Springer, 2016, pp. 53–70.

[46] C. S. R. Stoean and M. Preuss, "Approximating the number of attraction basins of a function by means of clustering and evolutionary algorithms," in *Proceedings of Research Notes in Artificial Intelligence and Digital Communications*. Reprograph Press, 2008, pp. 171–180.

[47] C. Stoean, M. Preuss, R. Stoean, and D. Dumitrescu, "EA-powered basin number estimation by means of preservation and exploration," in *Proceedings of International Conference on Parallel Problem Solving from Nature*. Springer, 2008, pp. 569–578.

[48] S. Maree, T. Alderliesten, D. Thierens, and P. A. Bosman, "Niching an estimation-of-distribution algorithm by hierarchical gaussian mixture learning," in *Proceedings of the Genetic and Evolutionary Computation Conference*, 2017, pp. 713–720.

[49] J.-W. Chen, M.-C. Lu, and T.-L. Yu, "Continuous optimization by hierarchical gaussian mixture with clustering embedded resource allocation," in *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion*, 2020, pp. 191–192.

[50] Z. Pitra, J. Repickỳ, and M. Holeňa, "Landscape analysis of gaussian process surrogates for the covariance matrix adaptation evolution strategy," in *Proceedings of the Genetic and Evolutionary Computation Conference*, 2019, pp. 691–699.

[51] A. Janković and C. Doerr, "Adaptive landscape analysis," in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 2019, pp. 2032–2035.

[52] P. Kerschke, M. Preuss, S. Wessing, and H. Trautmann, "Detecting funnel structures by means of exploratory landscape analysis," in *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, 2015, pp. 265–272.

[53] P. Kerschke and H. Trautmann, "The R-package FLACCO for exploratory landscape analysis with applications to multi-objective optimization problems," in *Proceedings of 2016 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2016, pp. 5262–5269.

[54] B. Bischl, S. Wessing, N. Bauer, K. Friedrichs, and C. Weihs, "MOI-MBO: multiobjective infill for parallel model-based optimization," in *Proceedings of International Conference on Learning and Intelligent Optimization*. Springer, 2014, pp. 173–186.

[55] S. Wessing and M. Preuss, "On multiobjective selection for multimodal optimization," *Computational Optimization and Applications*, vol. 63, no. 3, pp. 875–902, 2016.

[56] S. Wessing, M. Preuss, and G. Rudolph, "Niching by multiobjectivization with neighbor information: Trade-offs and benefits," in *Proceedings of 2013 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2013, pp. 103–110.