2009 Ninth International Conference on Intelligent Systems Design and Applications

Valley-Adaptive Clearing Scheme for Multimodal Optimization Evolutionary Search

Mostafa M. H. Ellabaan
School of Computer Engineering
Nanyang Technological University
most0001@ntu.edu.sg

Yew Soon Ong
School of Computer Engineering
Nanyang Technological University
asysong@ntu.edu.sg

Abstract

Recent studies [13, 18] have shown that clearing schemes are efficient multi-modal optimization methods. They efficiently reduce genetic drift which is the direct reason for premature convergence in genetic algorithms. However, clearing schemes assumed a landscape containing equal-spaced basins when using a fixed niche radius. Further, most clearing methods employ policies that favor elitists, thus affecting the explorative capabilities of the search. In this paper, we present a valley adaptive clearing scheme, aiming at adapting to non-uniform width of the valleys in the problem landscape. The framework of the algorithm involves hill-valley initialization, valley-adaptive clearing and archiving. Experimental results on benchmark functions are presented to demonstrate that the proposed scheme uncovers more local optima solutions and displays excellent robustness to varying niche radius than other clearing compeers.

1. Introduction

Most real world problems exhibit the property of having more than a local optimum solution. In particular, such problems are of great abundance in science and engineering [1-6]. Aerodynamic design, chemical isomers, scheduling and assignment problems represent some of the areas that possess fitness landscapes that are multi-modal in nature. Among these problems, most effort to date has concentrated on revealing the location and properties of the global optimum(s). Significantly lesser effort has been placed on identifying the set of all acceptable solutions. A motivating example for us in chemistry and physics is the discovery of low energy stable and meta-stable molecular structures which has remained an important and unsolved problem. Global optimization are often conducted at the expense of locating other low-lying isomers or otherwise known as local minima in the context of optimization. It is worth noting that isomers not only provide key insights into resultant properties but a statistical comparison between isomers also represents a more robust methodology for comparing models and determining fit to the quantum mechanical calculations [7,8].

In this paper, we are interested not only in finding one or more global optima but in identifying the set of all acceptable solutions. Such paradigms are commonly known as multi-modal optimization problems (MMOP). MMOPs can be classified into two main categories, according to the number and distribution of the local optimum set, namely, 1) finite MMOPs: a discrete set with finite optima representing the set of acceptable solutions, 2) infinite MMOPs where optima can be, for example, materialized in a circular manner similar to the waves of a lake when a stone is thrown in it [8]. In this work, our research interest is to deal with MMOPs containing a finite set of acceptable solutions.

In light of their excellent adaptabilities, evolutionary algorithms are well-established as strong candidates for handling multimodal optimization problems. One of their key strengths lies in their ability to maintain a diverse set of solutions in a single population [9-13] when appropriately designed. Hence, several new challenges to evolutionary computation have arisen when dealing with problems plagued with multiple optima. How to assess when or if a local optimum has been detected and how to discover and identify unique local optima in the evolution process once it has been detected remain to be among some of the open issues. To date, several efforts by researchers have been made to tackle such problems. The most common methods such as fitness sharing and clearing methods are summarized and compared in [13]. The comparison has shown that clearing methods efficiently reduce genetic drift and maintain multiple solutions. However, clearing methods generally assume the landscape to contain equal-sized basins, with a basin size larger than or equal to the clearing radius. Such information requires a prior knowledge about the landscape. An inappropriate value assumed would

978-0-7695-3872-3/09 $26.00 © 2009 IEEE
DOI 10.1109/ISDA.2009.115

seriously affect eventually the search performance. Further, some individuals in the clearing methods have no chance to participate in the reproduction operators, limiting the exploration capabilities of the clearing methods.

In this paper, we present a novel multi-modal optimization algorithm. In this algorithm, the initial population of individuals is sampled to sit in unique valley or basin of attraction. Individuals then undergo the reproduction operators. Individuals falling on a same valley are grouped together and subsequently categorized into elites and inferiors according to their fitness. Elites are enforced to survive to the next generation, while the inferiors are locally optimized. All unique elite solutions are then archived.

The paper is organized as follows: Section 2 provides a brief definition of the non-linear programming problem. Related work is presented in Section 3. The details of the proposed method are presented in section 4 while section 5 reports the results obtained from our computational study. The brief conclusion is then stated in Section 6.

2. Problem Statement

Here, we consider the general multi-modal nonlinear programming problem of the following form.

Minimize:

$$F(\mathbf{x}) \tag{1}$$

Subjected to

$$g_i(\mathbf{x}) \leq 0 \qquad i = 1,2,\dots,d$$

$$\mathbf{x_l} \leq \mathbf{x} \leq \mathbf{x_u}$$

where $\mathbf{x} \in \mathbf{R}^d$ is the vector of design variables, $\mathbf{x_l}$ and $\mathbf{x_u}$ are the lower and upper bound vectors, respectively. It is often the case that $F(\mathbf{x})$ and $g_i(\mathbf{x})$, which denote the objective and inequality functions, respectively, are computationally expensive to evaluate. A local optimum is a stationary point with vanished gradient and positive definite hessian matrix. These points can be mathematically expressed as $\mathbf{X_p} = \left\{\mathbf{x}_i | \left(\frac{dF(\mathbf{x_i})}{d\mathbf{x_i}} = 0\right) \ \& \ \left(\forall e_{i,j} \geq 0\right)\right\}$, where $\mathbf{x_i} \in \mathbf{R}^d$, d is the dimensional size, $\mathbf{X_p}$ is the finite set of all acceptable solutions, and $e_{i,j}$ is the eigenvalue of the hessian matrix $\mathbf{H_i}$.

3. Multi-Modal Optimization Methods: A Brief Review

Seeking for more than one solution is a common theme in numerous problems of science and engineering. Using genetic algorithm (GA) to address such problems is not generally recommended, since the conventional GA is designed to converge at single solution. In recent decades, many niching methods have been proposed to make genetic algorithms capable of tackling multi-modal optimization problems. The more popular approaches are fitness sharing [12], deterministic crowding [14], probabilistic crowding [15], clustering [16], restricted tournament selection [17], and the clearing methods [13, 18]. A comparison of these multi-modal optimization methods in [13] has reported that clearing methods are shown to be generally more efficient for exploration of the problem search spaces. Unlike other niching methods where resources are shared among similar individuals of the population, clearing methods only consider the elite individuals [13], while other less fit individuals are treated with a killer penalty. Petrowski [18] also reported that clearing methods effectively succeed in reducing the effect of genetic drift. Nevertheless, it is worth noting that clearing methods suffer from several limitations [13]. Particularly, inferior individuals have no chance of participating in the mating operations, limiting the capability of the algorithm in exploring the search space sufficiently well. There is also a poor usage of resources since those individuals that generate little effectiveness hog up the limited finite population slots. It is also worth noting as well that the efficacy of both canonical clearing [18] and modified clearing methods [13] remains to be sensitive to the configuration of $\mathrm{Q_{clear}}$ or clearing niche radius. An inappropriate value assumed would seriously affect the eventual search performance. Lastly, since clearing methods do not adapt to the nature of the problem landscape, there is a high chance that some local optimum solutions are left uncovered or missed in the search.

4. Valley Adaptive Clearing Evolutionary Search

Here, a valley-adaptive clearing multi-modal evolutionary search methodology is proposed. In this evolutionary search, the population of individuals is initiated with each individual falling in a different basin of attraction, using a *hill-valley detection scheme* (see section 4.1)[1]- A minimization problem is assumed. Each individual $\mathbf{x}$ is evaluated based on $F(\mathbf{x})$. Individuals then undergo selection, mutation and crossover. Thereafter, the yielding offspring undergoes the valley adaptive clearing scheme which

[1] A minimization problem is assumed

involves the *identification* (see section 4.2.1), *clearing* (see section 4.2.2) and *valley replacement* phases (see Section 4.2.3).

4.1 Hill-Valley Detection

The hill-valley detection procedure [19] begins by generating a line connecting two given points (**s** and **e**) in the Euclidean space. Subsequently, a number of intermediate points are sampled within the line. The fitness values of these points are then calculated. A valley existence, on one hand, is identified, if the fitness of any sampled points represents an improvement over that of the given points. Otherwise, a hill is found [20].

4.2 Valley-Adaptive Clearing Scheme

The valley-adaptive clearing scheme is composed of three core phases. The valley identification phase categorizes the population of individuals into groups of individuals sharing the same valley, denoted as $\mathbf{V}_{groups}$. Subsequently, the dominant individual (i.e., in terms of fitness value) of a valley group or $\mathbf{V}_{id}$ is archived if it represents a unique local optimum solution, while all other members of the same group undergo the valley replacement phase where relocation of these individuals to new basin of attractions or valleys are made so that unique local optimum solution elsewhere may be uncovered. In the event that no local optimum solution exists in a valley group, all individuals of the group will undergo the valley clearing stage where elite individuals are ensured to survive across the search generation while all others are relocated to new basin of the attractions.

4.2.1 Valley Identification Phase

The procedure of valley identification begins with the sorting of population individuals in ascending order according to fitness. Individuals are then grouped together if they share a common valley. Individuals belonging to the same valley group are then categorized according to their fitness into *elites* and *inferiors*. Elites are the fittest *k* individuals in a group, while the remaining individuals are the inferiors.

4.2.2 Valley Clearing Phase

Valley clearing is a process in which less fit individuals (or inferiors) are relocated out of the same basin of attraction, leaving valleys to be further exploited by the fittest individuals (elites). In the valley clearing process, each inferior member (**x**) of the valley group ($\mathbf{V}_{id}$) is relocated randomly in the range of Q_{clear}, i.e., the clearing niche radius, to 3 * Q_{clear}, whereas other individuals (or elites) are left unchanged for the purpose of exploiting the basin of attraction.

4.2.3 Valley Replacement Phase

The motivation behind valley replacement process is to reduce any computational resources wasted on rediscovering of valleys where the optima have already been uncovered. Individuals of the populations falling in previously encountered valleys are replaced with individuals in new basin of attractions, so as to bias the search towards previously unexplored region of the landscape.

4.3 Archiving Procedure

All optimum solutions found throughout the search are archived using an indexed database. Two data structures, the first is the array(s) or list(s) of discovered solutions, while the second is a hierarchical index or tree. Its nodes represent all cluster centers of the solutions found throughout the search, organized in a hierarchical manner according to the spatial order between solutions. The lists of solutions lie at the leave nodes of the index tree.

To keep the archive free of duplicates, we proposed a hybrid archiving procedure that combines a distance metric with hill-valley detection procedure to detect duplicates in the archive. In this procedure, the hill-valley detection procedure is employed only on selected archived optima that fall within a predefined distance of an optimum.

5. Empirical Study

In this section, we study the efficacy of the valley adaptive clearing scheme (AVAC) in comparison with two other existing clearing methods, particularly the canonical clearing (C) and modified clearing (MC) schemes, using several multi-modal benchmark test problems. The test problems considered in the study and experimental results obtained are presented in Sections 5.1 and 5.2, respectively.

5.1 Benchmark Test Problems

5.1.1 Two-Dimensional (2D) Test problems

Problem 1. The 2D Rastrigin Function

The 2D Rastrigin function Eqn. (2) is a typical multi-modal benchmark test problem used in evolutionary computation research. In the region specified below, the 2D Rastrigin function landscape contains 25 basins of almost equal size.

$$F(x,y) = 20 + x^2 + y^2 - 10(\cos 2\pi x + \cos 2\pi y) \quad (2)$$

where $x, y \in [-2.5, 2.5]$

Problem 2. The 2D Sines function

The 2D Sines function Eqn. (3) is a symmetric multi-modal test function with almost equal-size 49 basins.

$$F(x,y) = 1 + sin^2(x) + sin^2(y) - 0.1e^{-x^2-y^2} \quad (3)$$

Where $x, y \in [-10, 10]$

Problem 3. The 2D Multi-function

The 2D Multi-function Eqn. (4) is a challenging test case for multi-modal optimization. Its landscape includes 64 basins of varying sizes and heights.

$$F(x,y) = -1 - x\sin(4\pi\, x) + \; y\sin(4\,\pi\, y + \pi) \quad (4)$$

where $x, y \in [-2, 2]$.

5.1.2 20 Dimensional (20D) Test Problems

Problem 4. Hump Test problems

Hump function is multi-variable function, in which l basins are randomly located with different shapes and sizes such that the distance between two optima (**z** and **d**) are greater than ($r_{\mathbf{z}} + r_{\mathbf{d}}$), where $r_{\mathbf{z}}$, and $r_{\mathbf{d}}$ represent the radius of the basins around optimum **z** and **d**, respectively.

$$F(\mathbf{x}) = \begin{cases} -h_{\mathbf{k}}\left[1 - \left(\frac{d(\mathbf{x},\mathbf{k})}{r_{\mathbf{k}}}\right)^{\alpha_{\mathbf{k}}}\right] & if\ d(\mathbf{x},\mathbf{k}) \le \ r_{\mathbf{k}} \\ 0 & Otherwise \end{cases} \quad (5)$$

$\mathbf{x} \in \mathbf{R}^d, x_i \in [0,1]$, $h_{\mathbf{k}}$, α_k and $r_{\mathbf{k}}$ donating the depth, a factor that determine the basin shape, size of the basin around an optimum (k) are configured as 100, 1 and 1, respectively. number of basins is configured at 50. $d(\mathbf{x},\mathbf{k})$ is the Euclidean distance between an individual **x** and an optimum **k**.

5.2 Experimental Results

In this section, the experimental settings and results are reported. All clearing schemes are implemented in MATLAB development environment and experiments executed on a PC with 2.66 GHz Intel Duo Core CPU and 3 GB RAM. 50 independent runs are therefore executed to test the performance of the proposed scheme using the following algorithmic configurations: uniform mutation probability of 20%; scatter crossover probability of 60%, 20% of the population represents the elites and stochastic uniform sampling based selection. The number of sample points (m) in the hill-valley detection, the number of elites (k), Q_{clear} and the archiving-distance threshold are configured at 5, 2, 0.25 and 0.5, respectively. A population size of 50 and 150 is considered for the 2D and 20D problem, respectively. All the search terminates at a maximum generations of 100. In this section, performance analysis of clearing schemes with and without archive as well as the sensitivity of the archiving clearing schemes to the niche radius change are investigated in section 5.2.1 and 5.2.2, respectively

5.2.1 Performance Analysis of Clearing Schemes

We study the search performance of the clearing schemes with and without archiving in terms of percentages of uncovered optima and execution time. Experiments were, therefore, executed on different landscapes of varying degree of complexity and dimensionality. the performance of the clearing schemes (C-Canonical Clearing, and MC -Modified Clearing) and archiving clearing schemes (AC-Archiving Canonical Clearing, AMC – Archiving Modified Clearing, and AVAC- Valley-Adaptive Clearing) on all benchmark problems stated in section 5.1 is summarized in figures 1-4. Figures (1 and 3) show the percentages of optima uncovered by different clearing schemes on the 2D and 20D benchmark problems. The plotted results show that the archiving clearing schemes have rates of improvement ranging from 2-10 times over non-archiving schemes in the percentage of uncovered optima. It is also shown that the valley adaptive clearing scheme maintains the highest percentages of uncovered local optima on each of the different benchmark problems. The percentages of optima adaptive uncovered by AVAC, for example, vary in the ranges of 90-100%, whereas those uncovered by others clearing compeers vary largely in the range of 10-95%, respectively. Due to the large search space and dimensionality, the gaps between the percentages of the proposed scheme and other clearing schemes are observed to be very wide on the 2D Sines and 20D Hump function benchmark problems. Figures (2 and 4) show the execution time of different clearing schemes on the 2D and 20D benchmark problems.

The archiving clearing schemes execution time tends to be reasonably higher than those of non-archiving clearing schemes. However, in complex optimization where function evaluations are computationally expensive [22], the archiving overhead may be regarded as negligible. In addition, AVAC generally outperforms AMC. In particular, AVAC saves 30-60% of the execution time required by AMC. Since the archiving clearing schemes outperformed non-archiving clearing schemes in maintaining multiple optima, we will consider only the archiving clearing schemes for further experimental analysis.

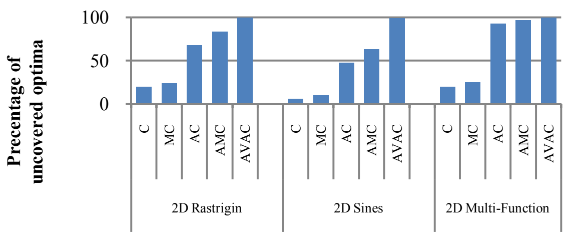

Figure 1: Percentages of uncovered optima by clearing schemes, on the three 2D benchmark problems.

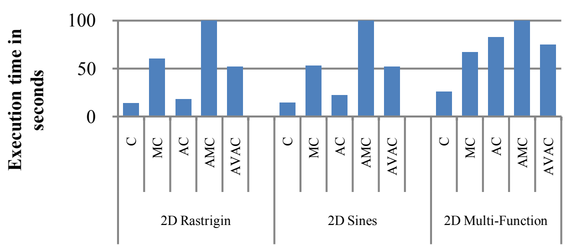

Figure 2: Execution time of different clearing schemes, on the three 2D benchmark problems.

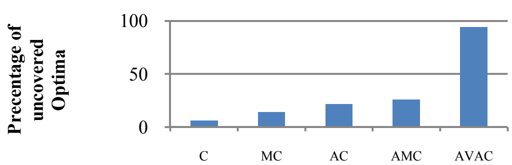

Figure 3: Percentages of uncovered optima by clearing schemes, on the 20D Hump Function.

5.2.2 Sensitivity Analysis of Niche Radius

Next, we study the sensitivity of the niche radius or Q_{clear} on the search performance of the archiving clearing schemes. In figures 6, the percentages of local optima uncovered by the respective archiving clearing schemes, for different configurations of niche radius Q_{clear}, on the four benchmark problems are summarized to vary largely in the range of 10-100%, on the four benchmarks. In contrast, the valley adaptive clearing scheme fares significantly better in the range of 85-100%. The percentage standard deviations of local optima uncovered by AVAC, AC and AMC vary in the ranges of 1.5-5.6, 5.5-20 and 7-23.5, respectively, demonstrated high sensitivity of the latter two methods to $Q_{clear,}$ the niche radius.

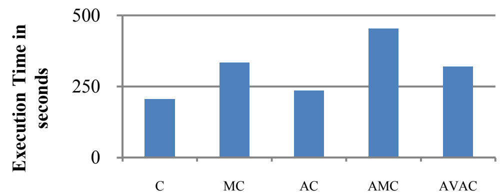

Figure 4: Execution time of different clearing schemes on 20D Hump Function.

6. Conclusion

In this paper, we proposed a new multi-modal optimization evolutionary search scheme -the valley adaptive clearing scheme- involving hill-valley initialization, valley-adaptive clearing and archiving. In this study, we also considered incorporating the archiving procedure with the existing clearing scheme. Performance of the clearing schemes with and without archiving was compared. The results showed that the archiving clearing schemes generally maintained much more optima than existing non-archiving clearing schemes. Among the archiving clearing schemes, experimental results showed that the proposed valley-adaptive clearing scheme maintained the highest percentage of uncovered optima with varying radii and landscapes of different degrees of complexity and dimensionality. These results highlighted the adaptability and robustness of the proposed scheme.

REFERENCES

[1] F. Neri, J. Toivanen, G. L. Cascella and Y. S. Ong, "An Adaptive Multimeme Algorithm for Designing HIV Multidrug Therapies”, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 4(2), pp. 264-278, 2007.

[2] Y. S. Ong, P.B. Nair and A.J. Keane, “Evolutionary Optimization of Computationally Expensive Problems via Surrogate Modeling”**,** American Institute of Aeronautics and Astronautics Journal, 41(4) pp. 687-696, 2003

[3] Y. S. Ong, P. B. Nair and K. Y. Lum, “Max-Min Surrogate-Assisted Evolutionary Algorithm for Robust

Design", IEEE Transactions on Evolutionary Computation, 10 (4), pp. 392-404, 2006.

[4] Z. Zhu, Y. S. Ong and M. Dash, "Markov Blanket-Embedded Genetic Algorithm for Gene Selection", Pattern Recognition, 40 (11) pp. 3236-3248, 2007.

[5] Y. S. Ong, K. Y. Lum and P. B. Nair, "Evolutionary Algorithm with Hermite Radial Basis Function Interpolants for Computationally Expensive Adjoint Solvers", Computational Optimization and Applications, 39(1) pp. 97-119, 2008.

[6] S. Hasan, R. Sarker, D. Essam and D. Cornforth, 'Memetic algorithms for solving job-shop scheduling problems', Memetic Computing, 1(1) pp. 69-83, 2009.

[7] Q. C. Nguyen, Y. S. Ong, H. Soh and Jer-Lai Kuo, "Multiscale Approach to Explore the Potential Energy Surface of Water Clusters $(H2O)_8$ n<=8", Journal of Phys. Chem. A, 112 (28).

[8] X. F. Fan, Z. Zhu, Y. S. Ong, Y. M. Lu, Z. X. Shen, and Jer-Lai Kuo, "A Direct First Principle Study on the Structure and Electronic Properties of BexZn1-xO", Applied Physics Letter 91, 121121, 2007.

[9] X. Yu, K. Tang, T. Chen, and X. Yao. Empirical analysis of evolutionary algorithms with immigrants schemes for dynamic optimization, Memetic Computing 1(1), pp. 3-24, 2009.

[10] J. Barcardit, E. Burke and N. Krasnogor, 'Improving the scalability of rule-based evolutionary learning', Memetic Computing 1(1) pp.55–67, 2009.

[11] R. Lung, and D. Dumitrescu, "A New Evolutionary Modal for Detecting Multiple Optima", in GECCOO 2007, London, UK. p. 1296 – 1303.

[12] D. Goldberg, and J. Richardson, "Genetic algorithms with sharing for multi-modal functions optimization. in Genetic Algorithms and Their Applications", in ICGA-87 pp. 44–49, 1987.

[13] G. Singh, and K. Deb, "Comparison of Multi-Modal Optimization Algorithms based on Evolutionary Algorithms", in GECCO, Seattle, Washington, USA pp. 1305 – 1312, 2006.

[14] S. Mahfoud, "Niching Method for Genetic Algorithms", in Department of Computer Science, 1995, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[15] O. Mengsheol and D. Goldberg, "Probabilistic crowding: Deterministic crowding with probabilistic replacement", in GECCO, pages 409–416, 1999.

[16] X. Yin and N. Germay. "A Fast Genetic Algorithm with Sharing Scheme Using Cluster Analysis Methods in Multi-modal Function Optimization. In Proceedings of the International Conference on Artificial Neural Nets and Genetic Algorithms, p.450–457, 1993.

[17] G. Harick, "Finding Multi-Modal Solutions Using Restricted Tournament Selection", in ICGA. p. 24–31, 1997.

[18] A. Petrowski, "A Clearing Procedure as a Niching Method for Genetic Algorithms, in ICEC, 1996.

[19] J. Branke, "Evolutionary Optimization in Dynamic Environments", 2002, Kluwer Academic Publishers.

[20] J.W. Gan, K. Warwick, W. Reading, "Dynamic Niche Clustering: a fuzzy variable radius niching technique for multimodal optimisation in GAs", in Congress on Evolutionary Computation. 2001. p. 215-222.

[21] D. Zaharie, "A Multipopulation Differential Evolution Algorithm for Multimodal Optimization", in R. Matousek, P. Osmera (eds.), Proc. of Mendel 2004, 10th International Conference on Soft Computing, Brno, June 2004, pp. 17-22.

[22] Z. Z. Zhou, Y. S. Ong, M. H. Lim and B. S. Lee, "Memetic Algorithm using Multi-Surrogates for Computationally Expensive Optimization Problems", Soft Computing Journal, Vol. 11, No. 10, pp. 957-971, August 2007.

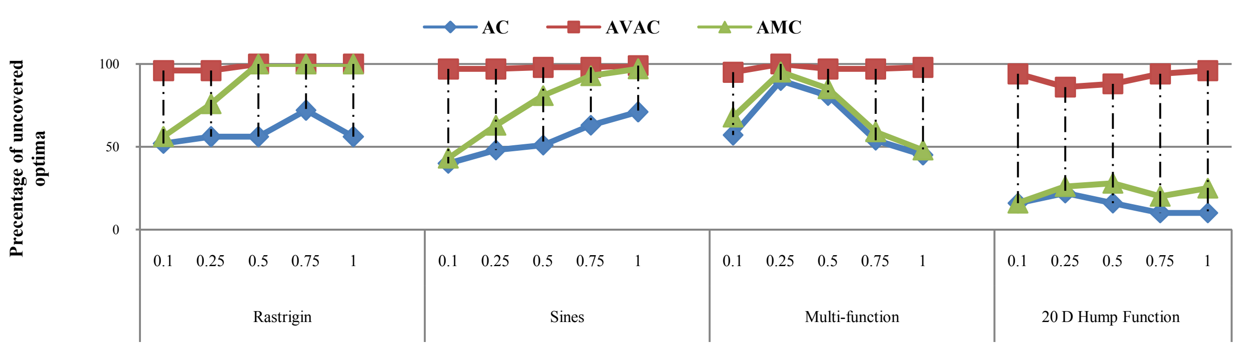

Figure 5: Percentages of optima uncovered by archiving clearing schemes on the benchmark problems, for different niche radii (Q_{clear}).