Handling Diversity in Evolutionary Multiobjective Optimistion

Nasreddine Hallam
nasreddine.hallam@nottingham.edu.my
Peter Blanchfield
peter.blanchfield@nottingham.edu.my
University of Notingham, Malaysia Campus

.
Graham Kendall
University of Notingham, Jubilee Campus.
gxk@cs.nott.ac.uk

Abstract- In Evolutionary Multiobjective Optimisation (EMO), the diversity of the set of non-dominated solutions used to be handled by the niching and fitness sharing technique. The main downside of this technique is the need to set the niche radius.
Quite recently, new techniques have emerged and proved to be more successful. The grid-based density of the Adaptive Grid Algorithm (AGA), the crowding-distance technique of the Nondominated Sorting Genetic Algorithm (NSGA-II), and the archive truncation procedure of the Strength Pareto Evolutionary Algorithm (SPEA2) are the latest successful methods that ensure a better diversity than the traditional less effective and computationally expensive niching method. In this work, a crowding-dispersion technique which is based on the Pareto Potential Regions (PPR), is proposed and compared to three recent techniques.

1 Introduction

The objectives in a Multiobjective Optimisation Problem (MOP) (e.g., cost, performance, quality of service, etc.) are often contradictory and/or conflicting with each other, and are supposed to be optimised (minimised and/or maximised) simultaneously. Techniques and algorithms that solve MOPs constitute a class of optimisation methods that has its root in the seminal work of Edgeworth and Pareto [Pareto 1896].

Mutli-Objective Evolutionary Algorithms (MOEAs), a metaphor of natural evolution, have established themselves as amongst the successful methods for solving MOPs. In fact, MOEAs had, interestingly through the past two decades, evolved and gone through a sophistication process and emerged as the *de facto* MOP solver. The most attractive feature of these algorithms is their capability to hunt for a set of solutions in a single run by considering a population of potential solutions. At the end of the execution, a MOEA is supposed to detect an effective set of solutions (known as the non-dominated set - usually an approximation of the Pareto set) which is presented to the Decision Maker (DM). The DM, in turn, will have to choose among the solutions of this non-dominated set, or may refine the original MOP so that the non-dominated set is reduced which will ease the process of decision making.

Therefore, a diverse set is primordial. The diversity brings crucial knowledge to the DM- the more diverse the solutions are, the better informed the DM is about the range and the spectrum of the possible solutions. This is because if the solutions in the non-dominated set are concentrated in one part of the search space (absence of diversity), the unexplored parts of this search space may well contain efficient solutions that might very well be of interest to the DM.

Traditionally, this diversity was maintained by using the *fitness sharing* [Goldbreg 89 ; Goldbreg and Richardson 1987 ; Horn 1997 ; Mahfoud 1995]. More recently, it has been shifted to density-based methods that work on a grid-like subdivided objective space [Knowles and Corne 2000, 2003 ; Lu 2002]. The AGA [Knowles and Corne 2003], a grid-based method, ensures (in a simpler and efficient way) the diversity of the set of solutions.

The crowding-distance, a slightly better method than AGA with regard to diversity, was proposed in NSGA-II [Deb et al 2002]. The archive truncation procedure, a more optimal technique but computationally expensive, has been proposed in SPEA2 [Zitzler et al. 2001].

The aim of this study is to present a new method and to compare the different diversity preserving methods.

In the next section, the theoretical background is reviewed. Section 3 presents the most popular diversity-handling techniques. Their shortcomings are also discussed. Section 4 is devoted to a new archiving approach (PPR), where a specific technique from this approach that deals with diversity (called crowding-dispersion) is presented.

The empirical study, (experimental methodology, results, and inferences), is presented in section 5. Finally the last section concludes this work.

2 Theoretical Foundation

An optimisation problem includes a set of decision variables or design variables, a set of objective functions, and a set of constraints.

The set of decision variables are those parameters the designer might tune in order to adjust or optimise the overall performance of a given system. The set of objective functions are those mathematical expressions that combine the set of decision variables in, possibly, different ways, each denoting a characteristic of the system. The set of constraints delimit the space of the decision variables.

In the following, we assume, without loss of generality, a minimisation problem.

2.1 Concepts and Definitions

Let Ω be the space of the decision variables, F the set of objective functions, and C the set of constraints.

The objective and constraint functions are functions of the decision variable x in Ω.

Let y be the image of x in the objective space induced by F. Let Λ, called the objective space, be the image of Ω by F.

Definition 1 *MOP's purpose is:*

$$\begin{cases} \textit{minimise} & F(x) = (f_1(x)\ , ..., f_k(x)) \in \Lambda. \\ \textit{Subject to} & C = (c_1(x)\ ,\ c_2(x), ..., c_l(x)) \leq 0 \\ \textit{Such that} & x = (x_1, x_2\ ,\ ...\ ,\ x_n) \in \Omega. \end{cases} \tag{1}$$

where $f_i(x),\ i = 1..k$, are the k objective functions, $k \geq 2$, and $x = (x_1, ..., x_n)$ is the n-dimensional decision vector, for which each x_i will be optimised.

Let $X = \{x \in \Omega,\ C(x) \leq 0\}$ denotes the feasible decision space. Similarly, let $Y = F(X) = \{y \in \Lambda\ ,\ C(F^{-1}(y)) \leq 0\}$ be the feasible objective space.

2.1.1 To Dominate or not to dominate

Let $u = (u_1, u_2, ..., u_k)$ and $v = (v_1, v_2, ..., v_k)$ be two k-dimensional objective vectors in R^k.

If compared to each other, one of the following situations may apply:

$$\begin{cases} u = v & \text{iff} & u_i = v_i\ \forall i = 1..k \\ u \preceq v & \text{iff} & u_i \leq v_i\ \forall i = 1..k \\ u \prec v & \text{iff} & (u \preceq v) \wedge (u \neq v) \\ u \sim v & \text{iff} & (u \not\preceq v) \wedge (v \not\preceq u) \end{cases} \tag{2}$$

2.1.2 Pareto-optimal Set

The Pareto-optimal Set, P^*, is constituted of those decision variables whose corresponding objective vectors are not dominated by any other vector in the objective space.

$$P^* = \{x_1 \in X |\ \nexists\, x_2 \in X\ s.t\ y_1 = F(x_1)\ ,\ y_2 = F(x_2)\ and\ (y_2 \prec y_1)\} \tag{3}$$

These vectors are also known as efficient, non-admissible, or non-inferior solutions.

2.1.3 Pareto-optimal Front

The Pareto-optimal front, PF^*, is the image of the Pareto-optimal set P^*. PF^* contains all those objective vectors that are not dominated by any vector in the objective space.

$$PF^* = \{y \in Y\ |\nexists y' \in Y\ s.t\ y' \prec y\} \tag{4}$$

Members of PF^* are often termed non-dominated objective vectors.

2.1.4 Non Dominated Set Filter

For each set of objective vectors $A \in Y$, a special subset can be derived by filtering non dominated objective vectors.

$$ND(A) = \{a \in A\ |\quad \nexists b \in A,\ b \prec a\}. \tag{5}$$

The Pareto Front PF^* is exactly $ND(Y)$.

2.1.5 Local Pareto Front

Given a population at iteration t, Pop^t, its corresponding Local Pareto Front, PF^t, consists of those non-dominated objective vectors, i.e., $PF^t = ND(Pop^t)$.

2.1.6 Archive

An (unbounded) archive at iteration t, denoted by A^t, contains all the non-dominated objective vectors found so far, up to iteration t.

$$A^t \subseteq Y\ ,\ A^t = ND(\bigcup_{i=1}^{t} PF^i)$$

However, in practice, an archive is of limited size and therefore, a procedure that selects which extra non-dominated objective vectors must be chosen for deletion .

3 Density Handling in MOEA

Diversity has been traditionally implemented by techniques based on Pareto Niching and Fitness Sharing ([Goldberg 1989], [Horn 1997]).

In the last five years, more efficient techniques have emerged. [Knowles and Corne 2003] developed a grid-based density technique that is more attractive than Niching and Sharing technique. One of the improvement brought to NSGA [Srinivas and Deb 1995] was the proposal of a new technique called *crowding-distance*. Similarly, the earlier SPEA [Zitzler and Thiele 1999] has also been upgraded by fixing, among other problems, the diversity issue.

In this section, these techniques are briefly explained and some of the noticeable shortcomings are highlighted. A more detailed discussion can be found in [Hallam 2004].

3.1 Pareto Niching and Fitness Sharing

A premature convergence is the Achille's heel of an 'elitist' evolutionary algorithm. To curb such drawbacks, Goldberg introduced a new technique that favours unexplored regions of the search space by penalising the crowded regions and rewarding the scarce regions.

Those individuals that are crowded in a region, hence the term niche, will each have its original fitness reduced by dividing it by the number of neighbours in the same niche. In this way, those individuals that are isolated have a better chance of being selected in future generations since their fitness has not been reduced, and consequently exploration is set in this isolated region. The shared fitness of a solution i is its old fitness divided by the niche count:

$$F(i) = \frac{F^{old}(i)}{\sum_{j \in Pop} s(d(i,j))}$$

The sharing function $s(d(i,j))$ is zero if the distance of a j solution to the i solution is bigger than the niche radius (σ_{share}), otherwise a value in $]0, 1]$ is returned.

While the technique of Niching and Fitness Sharing is used for maintaining multiple optima in a SOP, it has been

used in MOP to diversify the set of Pareto solutions. As stated by [Horn 1997] and other researchers ([Fonseca and Fleming 1995],[Coello 1999], etc...), the main difficulty with this technique is the setting of the parameter σ_{share}. How large the niche should be? i.e., what is the best value for the niche radius? To answer this question, practitioners often conduct a set of test experiments just to determine the value for the niche radius, after which, they continue with their main experiments.

3.2 The Adaptive Grid Approach: Analysis and Shortcomings

Recently, [Knowles and Corne 2003] have developed an alternative technique, Archiving Grid Algorithms, that is based on subdividing the objective space into equal polytopes that constitute a hyper grid. Each polytope can be seen as a region within which lay objective vector solutions. From a computational point of view, this technique is attractive in the sense that the grid can be implemented as a k-dimension array matrix, where each array element represents a region and the value in each array element represents the number of objective vectors that reside in the corresponding region, i.e. the density of the region.

In this way, it would be much easier and more efficient to direct the search process to explore less crowded regions thus promoting diversity. It is worth noting that this technique has also been adopted by other researchers (Lu 2002; Coello et al 2004). However, serious drawbacks have recently been discovered (see [Hallam 2004]).

3.2.1 The Adaptive Grid Diversity-Preserving Technique: How It Works

As mentioned earlier, the objective space is subdivided into equal regions (polytopes). The vector solutions, generated over time, that are non-dominated are admitted into the archive as long as the size of the archive is less than the maximum size allowed, $arcsize$. Each region, however, updates its count any time a vector is inserted into it (or removed).

The interesting situation, which occurs quite frequently, is whenever the archive becomes full and a new non-dominated solution is discovered. In this case, the new vector solution will be inserted in the archive replacing one that is located in the most crowded (dense) region. This replacement greatly increases the well-spread distribution of the solutions throughout the Pareto front.

For instance, consider the illustrative example depicted in Figure 1.

The $arcsize$ is set to 8 and the archive is already full. It can be seen that the most crowded region is $R_{2,3} = \{b, c, d\}$. Therefore once the new (non-dominated) vector solution g is generated, one random vector is chosen from region $R_{2,3}$to be removed from the archive and the new vector solution g is admitted.

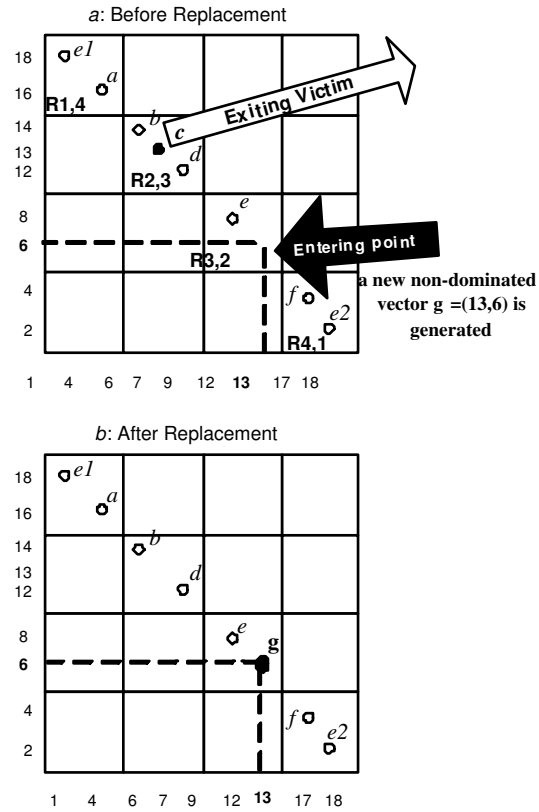

Figure 1: AGA based Diversity: An Example.

3.2.2 The Adaptive Grid Diversity-Preserving Technique: Shortcomings

The previous illustrative example ”exercised” to some extent the strength of the Adaptive Grid. However, a deeper look suggests that the judgement of a good spread is as effective as the judicious choice of the region size. It can further be argued that no matter how good the granularity of the grid is, it can always be the case that the most dense region is overlooked no matter what size of grid is used!

These shortcomings can be revealed through the simple illustrative example depicted in the following two figures, Figure 2 and Figure 3.

According to the simple definition of density (number of individuals in a given area). The top-left region $R_{1,4}$ of Figure 2 is the most dense region. Therefore if a new non-dominated vector candidate is generated, the exiting victim should be picked randomly from this crowded region $R_{1,4}$.

However, with exactly the same Pareto front, if the grid is further subdivided , for instance into an 8x8 regions (see Figure 2.b), the most crowded region is $R_{5,4}$ which is a subregion of the region $R_{3,2}$ of Figure2.a.

Notice that in Figure 2.a, region $R_{3,2}$ is totally different from the earlier supposed-to-be dense region $R_{1,4}$.

There was a gain of finer insight of the crowding and spreading of Pareto vectors with the increase of grid granularity. This means that we might actually overlook the most crowded region and as a result either the diversity is not well guaranteed or the spread velocity is slow.

Another more problematic situation arises when some

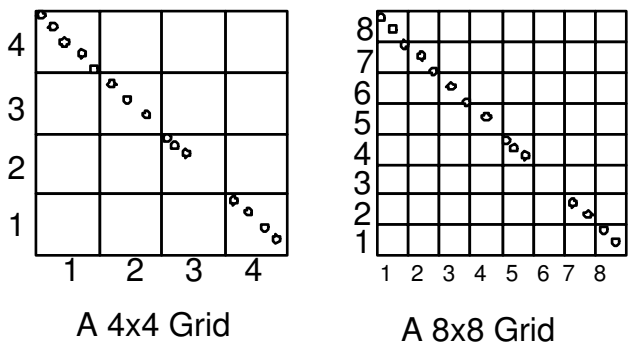

Figure 2: AGA based Diversity: Anolmaly 1.

Pareto points are crowded along the extreme common corners of two or more neighbouring regions. This interesting crowding, which might be the one sought for (the most crowded one), is simply overlooked no matter how many times the grid resolution is refined.

In Figure 3, none of the regions 1, 2, and 3 are the most crowded since they contain three vector solutions each. However, it can be seen that the non-apparent region between R2 and R3 harbour the most crowded subset of the solution vectors. Hence, the concept of crowding/density is just overstepped.

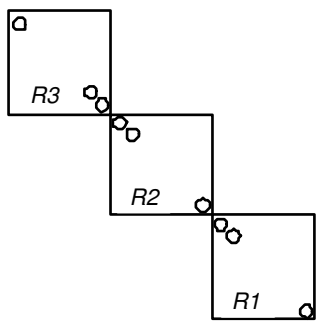

Figure 3: AGA based Diversity: Anomaly 2.

3.3 Crowding-distance technique (NSGA-II)

This technique was proposed in NSGA-II [Deb et al. 2002]. It sorts (in ascending order) the set of solutions according to each objective function. The extreme solutions (the solution with the smallest value and the one with the highest value) are assigned an infinite distance value. The remaining solutions are assigned distance values equal to the absolute difference of the objective values of two adjacent solutions. The same process is repeated for all other objectives, and the distances are accumulated.

The crowding-distance of a solution is simply the cumulative sum of the solution distances corresponding to each objective. As example, solution i in Figure 4 is assigned a crowding-distance equal to $d_1 + d_2$.

Then the most crowded solution is the one with the smallest crowding-distance value. Therefore in the case where the size of a set of non-dominated solutions exceeds the size of the archive, the right solutions to remove from the set are the ones with the smallest values of the crowding-distance.

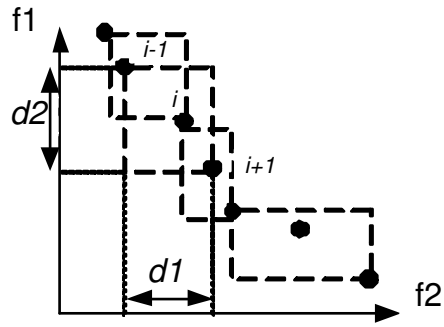

Figure 4: an Example of Crowding-Distance in NSGA2.

The main problem with this technique can be illustrated in Figure 5. Assume that we can keep at most 5 individuals in our population, and after the addition of the point np into the temporary population, the point a seems to be wrapped inside the smallest cuboid. Since all the points have the same level of dominance, the point a, which has the smallest crowding distance, would be removed.

A closer look would suggest that either c or d should be removed in order to have a better spacing.

A point in a large (or worse the largest) cuboid (but very close to one of its neighbour) can be the actual cause of a bad diversity than a point (equidistantly positioned) in the smallest cuboid. This fact is a key factor in the simulation results reported later in section 5.

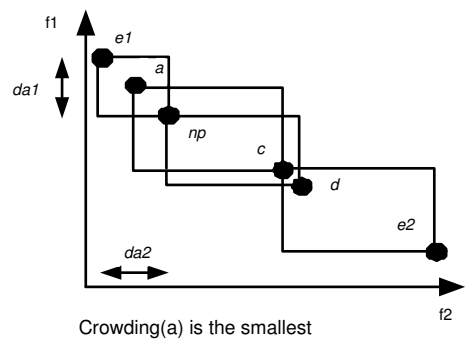

Figure 5: NSGA2 Crowding-Distance Truncation.

3.4 Archive Truncation Procedure (SPEA2)

In this technique, the archive for the next generation is referred to as P'_{t+1}, and its size $|P'_{t+1}| = N$ is fixed. Non-dominated solutions are copied into the archive. Three cases may arise. If the non-dominated solutions exactly fit the archive, then the process of filling the archive is terminated. If the non-dominated set is smaller than the archive, then the remaining slots of the archive are filled with the best solutions from the rest of the population.

The interesting case is when the number of non-dominated solutions exceeds the size of the archive. In this

case some non-dominated solutions must be removed. This is done by invoking the *archive truncation procedure* which iteratively removes one solution at a time until the size of $|P'_{t+1}| = N$. The individual i to be removed is the one that, $\forall j \in P'_{t+1}, i \leq_d j$.

$$\begin{aligned} i \leq_d j :\Leftrightarrow \forall\ 0 < k < |P'_{t+1}| : \sigma_i^k = \sigma_j^k \quad \vee \\ \exists\ 0 < k < |P'_{t+1}| : \\ [(\forall\ 0 < l < k : \sigma_i^l = \sigma_j^l) \wedge \sigma_i^k < \sigma_j^k] \end{aligned} \tag{6}$$

where σ_i^k denotes the distance of i to its k-th nearest neighbour in P'_{t+1}.

The main difficulty with this method is the need to compute all the inter-distances between all the individuals of the archive. This will later be seen in the experimental section.

4 PPR: Potential Pareto Regions

These are dynamic regions within which any generated vector solution is automatically non-dominated with regard to all the current non-dominated solutions. The non-dominated set is sorted according to one objective and each two immediate neighbours delimit one PPR (see Figure 6). The idea is to maintain a set of dynamic regions (PPRs).

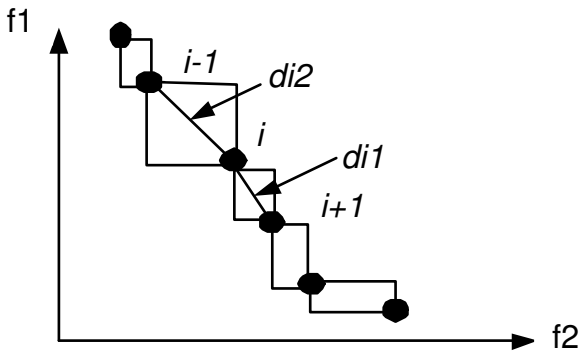

Figure 6: An Example of a PPR.

Definition 2 *A PPR is a hyper-area delimited by two immediate neighbouring non-dominated points of the archive.*

Let M^t be the archive, i.e., the set of current Pareto solutions.

Let $z^{t+1} = Gen(t+1)$ be a solution generated by the function Gen at iteration $t+1$.

Let PPR^{xy} be a Potential Pareto Region delimited by two (neighbouring) non-dominated solutions x and y.

4.1 How to construct a 2-D PPR

Arrange the $z^i, i = 1, |M^t| - 1$, such that $z^{i,k} > z^{i+1,k}$ for a given k. Without loss of generality we set k to 1.

Definition 3 *SPPR is the set of the PPRs ordered according to "neighbourhood" order.*
$SPPR = \{PPR^{z^i z^{i+1}} \mid z^{i,k} > z^{i+1,k}\ ,\ i = 1..|M^t| - 1\}$

Where $PPR_{ub_{1,t}}^{z^i z^{i+1}} = z^{i,1}$, $PPR_{lb_{1,t}}^{z^i z^{i+1}} = z^{i+1,1}$, $PPR_{ub_{2,t}}^{z^i z^{i+1}} = z^{i+1,2}$, and $PPR_{lb_{2,t}}^{z^i z^{i+1}} = z^{i,2}$, and $ub_{k,t}$, $lb_{k,t}$ respectively denote the upper and lower bounds of the k-th objective.
/* 1 is the chosen reference dimension, 2 is the other dimension.*/

Note that at any point in time, the size of the list of PPRs, $|SPPR|$, is always equal to the size of the archive minus 1. In Other words, $|SPPR| = |M^t| - 1$.

Lemma 1 *If there exists a PPR which contains z^{t+1}, then z^{t+1} is incomparable to any point in the archive.*
More formally, $z^t \in PPR^{xy} \Rightarrow \forall a \in M^t\ ,\ z^{t+1} \sim a$.

Proof of Lemma 1 Assume $\exists\ PPR^{xy}\ , z^{t+1} \in PPR^{xy}$.

The purpose is to prove that $\forall a \in M^t\ ,\ z^{t+1} \sim a$.

Assume that $\exists a \in M^t$ and that a is NOT incomparable to z^{t+1}, i.e., $(z^{t+1} \succ a) \vee (z^{t+1} \prec a)$.

Assume without loss of generality that $y_1 < x_1$ and $x_2 < y_2$.

Now PPR^{xy} is defined by all points z, $(y_1 < z_1 < x_1)$ and $(x_2 < z_2 < y_2)$.

There are two cases as to whether a is a neighbour of x or a neighbour of y.

1 . Assume the point a is a neighbour of y, $(a_1 < y_1)$ and $(y_2 < a_2)$.

   i) a dominates z^{t+1} means that $a_1 < z_1^t$ and $a_2 < z_2^t \rightarrow a_2 < z_2^t < y_2$, a contradiction. Therefore a cannot dominate z^{t+1}.

   ii) a is dominated by z^{t+1} means that $z_1^t < a_1$ and $z_2^t a_2 \rightarrow y_1 < z_1^t < a_1$, a contradiction. Therefore z^{t+1} cannot dominate a.

2 . Assume the point a is a neighbour of x, $(x_1 < a_1)$ and $(a_2 < x_2)$.

   i) a dominates z^{t+1} means that $a_1 < z_1^t$ and $a_2 < z_2^t \rightarrow a_1 < z_1^t < x_1$, a contradiction. Therefore a cannot dominate z^{t+1}.

   ii) a is dominated by z^{t+1} means that $z_1^t < a_1$ and $z_2^t a_2 \rightarrow x_2 < z_2^t < a_2$, a contradiction. Therefore z^{t+1} cannot dominate a.

4.2 Diversity With Two Indicators: Crowding and Dispersion

Given the current archive at time t, we compute the *degree* of crowding and the *extent* of distribution of each non-dominated solution. For the sake of an efficient computation, we maintain a matrix that stores, for each non-dominated vector, two distances (Euclidian distance) to its immediate neighbours. Recall that vectors in the archive are ordered according to one objective dimension and hence the immediate neighbourhood is based on this order.

If the size of the archive is n, there is a need to compute $n-1$ distances. Remember that the non-dominated vectors

Table 1: PPR

Efficient Solution	$Distance_{Neighbour1}$	$Distance_{Neighbour2}$
1 : first	-	$d_{first2nd}$
...	...	...
h	...	d_{hi}
i	d_{hi}	d_{ij}
j	d_{ij}	d_{jk}
...	...	...
n : last	d_{1last_last}	-

are (in the matrix) ordered according to one of the k objectives. Based on this matrix, the two important indicators, namely *crowding* and *dispersion* are computed as follows:

$$crowding(i) = min(d_{hi}, d_{ij}) \tag{7}$$

$$dispersion(i) = max(d_{hi}, d_{ij}) \tag{8}$$

It is obvious that the set $\{i, \min_{i \in M^t}(crowding(i))\}$ contains at least two elements. Assume that i is the vector with the minimum crowding value and that $d_{hi} > d_{ij}$, then the vector j, the neighbour of i, has also the minimum crowding. In this situation, which happens all the time, the most crowded among these two vectors is the one with the minimum dispersion. This is intuitively plausible since the closer the point to its immediate neighbours the more crowded the region encompassing all of them is.

4.3 How to find the most crowded vector solution

Simply get a vector solution *i* such that $crowding(i)$ is the minimum among the rest of the vector solutions (see Table 1). This will return at least two vector solutions (the neighbours). Then the one with the minimum dispersion value is selected as the most crowded one.

$$\begin{array}{l} IndexCrowded = min\ \ (Distance_{Neighbour2}(1..n-1)) \\ if\ IndexCrowded = 1\ then\ IndexCrowded++\ ; \\ else\ \ if\ Distance_{Neighbour2}[IndexCrowded-1] > \\ \qquad\qquad Distance_{Neighbour2}[IndexCrowded+1] \\ \qquad then\ \ IndexCrowded++\ ;\ endif \\ endif \end{array} \tag{9}$$

Note that the extremums are not included in the above process.

5 Empirical Experiment

5.1 Measuring The Diversity

For these experiments, we use the Schott Spacing Measure [Schott 1995]. Schott proposed a metric called *Spacing* that measures the range variance of neighbouring vectors in the archive.

$$S = \sqrt{\frac{1}{n-1}\sum_{i+1}^{n}(d - d_i)^2} \tag{10}$$

where $d_i = \min_j(|f_i^1 - f_j^1| + |f_i^2 - f_j^2|)$.
d: The mean of all d_i , $\frac{1}{n}\sum_{i=1}^{n}(d_i)$.
$n = |PF^t|$. i.e., the number of objective vectors in PF^t
.

$S = 0$ means that all members of the archive are equidistantly spaced. However, as noted by [Van Veldhuizen 1999], the landscape of some MOP can inherently be not uniformly spaced, or composed of two or more "distant" from-each-other Pareto front curves.

Nevertheless, our experiments are to compare few methods, therefore, this metric is reasonably justifiable, since we are not interested in how close the metric value is to zero, but on the values of this metric returned by each method. The lowest value would mean a better method.

5.2 Testing Methodology

An archive of 100 non-dominated points is randomly generated from. Then, 100 sets of 20 non-dominated points (2, 000 candidates) are randomly generated. Each point is non-dominated with regard to the archive. All points are taken from the real domain $[0, 100]^2$ with a precision of 4 decimal points.

The points in each set are presented one by one to the archive (which means 20 spacing values are computed for each set). For the purpose of plotting the results, we take the average spacing[1] from each set. Thus this amounts to hundred averages for the whole testing. We have conducted 100 experiments of this type (100 random archives with their 2000 non-dominated candidates). This means that 200, 000 spacing values have been computed. Since all the results were of the same pattern, only one experiment is plotted in Figure 7.

It must be noted that, although possible, there is no need to consider the different shapes of the fronts (convexity, discrete, non-uniformity of the solutions,...) since the random nature of these (100) archives by default ensures a generic treatment of this issue.

[1]The experimental data, their results, and the codes can be freely downloaded from http://sepang.nottingham.edu.my/~ hnasreddin/emo/diversity

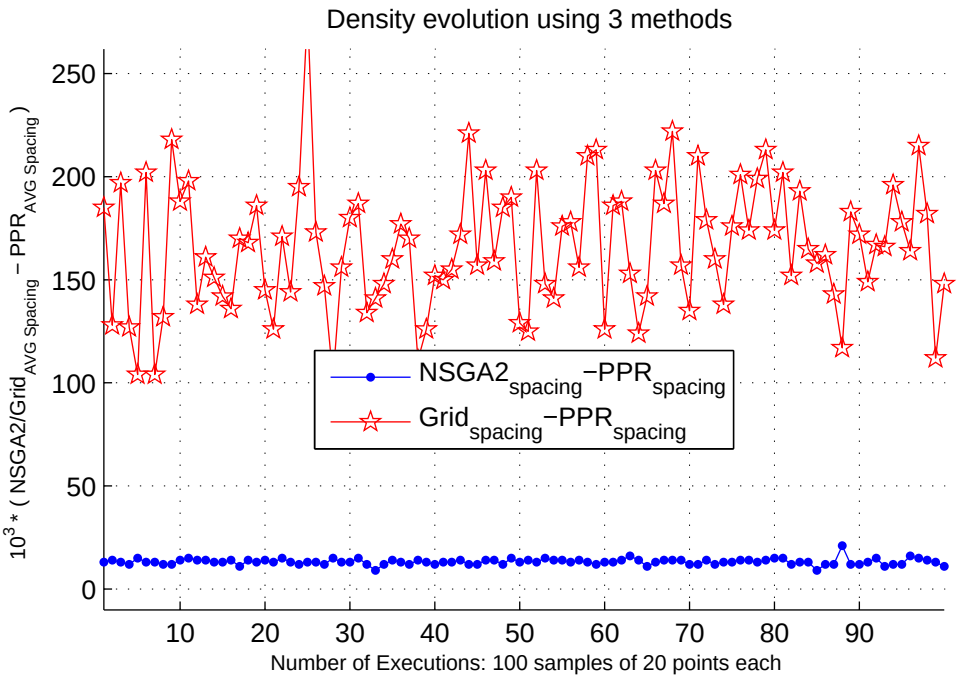

Figure 7: Density Handling Using PPR compared to NSGA2 and AGA

5.3 Results and Discussion

Figure 7 has been scaled up in order to clearly show the difference between the three methods. From that particular random test data, the difference between PPR and NSGA2 was less significant than the difference between PPR and AGA. However, in each of the 2000 attempts, PPR returned a better Spacing value.

The main reason why the NSGA2 crowding-distance technique has failed to better space between the non-dominated points can be traced back to its shortcoming in handling situations similar to the one illustrated in Figure 5, section 3.2.2. Points in small cuboid were chosen to be removed instead of other points in larger cuboids, but very much more closer to one of their neighbours. This is exactly what happened when we analysed the behaviour of NSGA2 crowding-distance based truncation against those of PPR and SPEA2.

Figure 8 shows that both the SPEA2 archive truncation procedure and the PPR diversity-preservation technique returned exactly (though not surprisingly) the same results. This can simply be explained that they are doing the same work.

However, according to Figure 9, PPR diversity-preservation technique is much faster. Both were executed on the same machine.

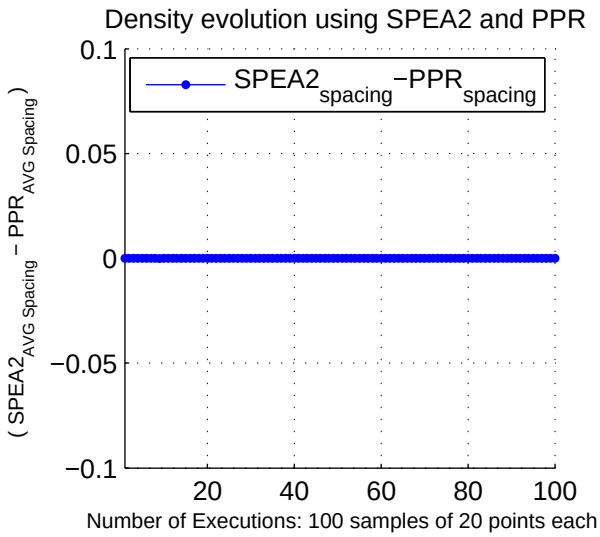

Figure 8: Density Handling Using PPR Compared to SPEA2

6 Conclusions

A new non-dominated objective vector solution would compete with the elements of the archive and not with the whole population. Therefore, it is trivial to build a list of neighbouring regions whose end-points are the non-dominated objective vectors themselves. As a result, the "factor-contribution-to-density" of a point is a matter of how large is the corresponding smallest region (cropwding), and eventually how evenly positioned this point is with regards to its corresponding two regions (dispersion). This is the basis of the PPR diversity preservation method.

We have also presented the state-of-the-art diversity preserving techniques used by the most popular MOEAs.

Some shortcomings inherent to the AGA and NSGA-2

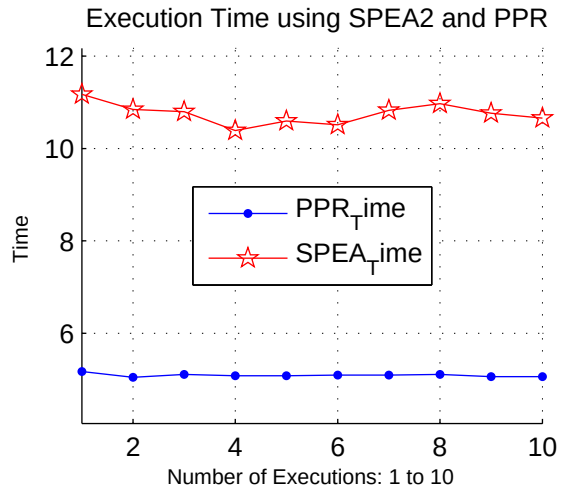

Figure 9: Execution Time- PPR Compared to SPEA2

diversity based techniques are pinpointed and examplified. Then simulation tests were conducted and confirmed these shortcomings.

The proposed method works as well as the SPEA2 archive truncation procedure. However it is much faster. Our future work will further investigate these two methods in order to construct a better method.

Bibliography

[coe99] C. A. Coello , "A Comprehensive Survey of Evolutionary-Based Multiobjective Optimization Techniques," *Knowledge and Information Systems*, vol 1(3), pp. 269-308, August 1999.

[coe04] C. A. Coello, G. T. Pulido and M.S. Lechuga, "Handling Multiple Objectives with Particle Swarm Optimization," *IEEE Transactions on Evolutionary Computation*, vol 8(3), pp. 256–279, June 2004.

[deb02a] K. Deb K, A. Pratap, S. Agarwal, and T. Meyarivan (2002a), "A fast and elitist multiobjective genetic algorithm: NSGA-II", *IEEE Transactions on Evolutionary Computation*, vol 6(2), pp.182-197.

[fon95a] C. M. Fonseca & P. J. Fleming, "An overview of evolutionary algorithms in multiobjective optimization,"*Evolutionary Computation*, vol 3(1), pp. 1-16. MIT Press Journals, 1995

[gol89] D. E. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning," Addison Wesley, 1989, Reading, Massachusetts.

[gol87] D. E. Goldberg, & J. Richardson, "Genetic algorithms with sharing for multi-modal function optimization," in textit Proc. 2nd Int. Conf. on Genetic Algorithms, 1987, pp.41-49, Lawrence Erlbaum.

[hal04] N. Hallam , "MultiObjective Evolutionary Algorithms: State of the Art", Technical Report, TR 1122004, The University of Nottingham, Malaysia Campus, Dec 2004.

[hor97] J. Horn, "Multicriterion decision making," In Bck, T., Fogel, D. & Michalewicz, Z. (eds.) Handbook of Evolutionary Computation 1: F1.9:1-F1.9:15. Oxford University Press, Oxford, England, 1997.

[kno00] J.D. Knowles & D.W. Corne (2000) "Approximating the non-dominated front using the Pareto Archived Evolution Strategy," *Evolutionary Computation*, vol 8(2), pp. 149-172, 2000.

[kno03] J.D. Knowles & D.W.Corne (2003) "Properties of an adaptive archiving algorithm for storing non-dominated vectors". *IEEE Transactions on Evolutionary Computation*, vol7(2), pp. 100-116.

[lu02] H. Lu, "State-Of-The-Art Multiobjective Evolutionary Algorithms-Pareto Panking, Density Estimation And Dynamic Population," PhD Thesis, Faculty of the Graduate College of the Oklahoma State University, August 2002.

[mah95] S. W. Mahfoud, "Niching methods for genetic algorithms," Ph.D. dissertation, 1995, University of Illinois, Urbana-Champaign.

[par1896] V. Pareto, "Cours d'Economie Politique, volume I et II," F. Rouge, Lausanne, Switzerland, 1896.

[sch95] J.R. Schott, "Fault Tolerant Design Using Single and Multicriteria Genetic Algorithm Optimization," Master's thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1995.

[sri94] N. Srinivas & K. Deb, "Multiobjective optimization using non-dominated sorting in genetic algorithms," *Evolutionary Computation*, vol 2(3), pp. 221-248. MIT Press Journals, 1994

[van99] D. Van Veldhuizen, "Multiobjective Evolutionary Algorithms: Classifications, Analyses, and New Innovations," PhD Thesis, Dept. of Electrical and Computer Eng., Graduate School of Engineering, Air Force Institute of Technology, Wright-Paterson AFB, Ohio.

[zit99] E. Zitzler & L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach," *IEEE Transactions on Evolutionary Computation*, vol 3(4), pp. 257-271, 1999.

[zit01] E. Zitzler, M. Laumanns, L. Thiele, "SPEA2: Improving the Strength Pareto Evolutionary Algorithm", *TIK Report Nr. 103, Computer Engineering and Networks Lab (TIK)*, Swiss Federal Institute of Technology (ETH) Zurich, May, 2001.