A Summary and Comparison of MOEA Algorithms

Daniel Kunkle

May 31, 2005

1 Algorithms Surveyed

The following MOEA algorithms are briefly summarized and compared:

- NPGA - Niched Pareto Genetic Algorithm (1994)
  - NPGA II (2001)
- NSGA - Non-dominated Sorting Genetic Algorithm (1994)
  - NSGA II (2000)
- SPEA - Strength Pareto Evolutionary Algorithm (1998)
  - SPEA2 (2001)
  - SPEA2+ (2004)
- PAES - Pareto Archived Evolution Strategy (2000)
- PESA - Pareto Envelope-based Selection Algorithm (2000)
  - PESA II (2001)
- MOPSO - Multi-Objective Particle Swarm Optimization (2003)
- ParEGO - Pareto Efficient Global Optimization (2005)
2 Considerations

The following are the main considerations when comparing MOEAs. How these considerations effect each of the algorithms discussed will be discussed when such information is available from the literature.

- Distance of the final approximation set to the true optimal set
- Diversity of Solution Set: diversity in either the decision variable space or the objective function space
- Convergence Time: the number of fitness evaluations required to ensure the algorithm has settled on a solution set.
- Ability to scale to high dimensionality of the decision variable space and objective function space
- Sensitivity to Noise: performance under noise fitness functions
- Sensitivity to Pareto-optimal front shape: convex, concave, discontinuous, etc.

3 NPGA - Niched Pareto Genetic Algorithm (1994)

- **Introduced** by Horn, Nafpliotis, and Goldberg in 1994 [15]

- **Diversity**: One of the first algorithms to directly address the diversity of the approximation set. All of the differences between NPGA and traditional GAs are localized in the selection mechanism. NPGA uses a modified tournament selection called *Pareto domination tournaments* along with *fitness sharing in the objective space* to maintain a diverse population.

- **Outdated**: NPGA has been shown to be inferior to most of the more recent MOEAs. For example, the comparison in [28] ranks NPGA fifth out of six MOEAs. Many other MOEAs use very similar techniques, including Pareto domination tournaments and fitness sharing. Examples are NPGA II, MOGA, and NSGA.
4 NPGA II (2001)

- **Introduced** by Erickson, Mayer, and Horn as an improvement to the original NPGA [11].

- **Main improvement** of NPGA II over NPGA is using the *degree of domination* (the number of solutions in the current population that dominate it) as the determining score in tournament selection. This method of deciding tournaments is deterministic, in contrast to the probabilistic method used by NPGA, which can result in a "noisier" search. See [12] for more information on the degree of domination score.

- **Limited Performance Comparisons**: in [11], the performance of NPGA II was only compared to a single objective optimization function and random search. More experimental results should be found before considering NPGA II superior to any other MOEA (including the original NPGA).

- **Sensitivity to parameters**: NPGA II, like NPGA, is very sensitive to the parameters controlling tournament selection and fitness sharing.

- Contrary to some sources, both NPGA and NPGA II use *continuously updated fitness sharing* to avoid chaotic perturbations of the population composition. See [22] for more discussion of the need for continuously updated fitness sharing when using tournament selection.

5 NSGA - Non-dominated Sorting Genetic Algorithm (1994)

- **Introduced** by Srinivas and Deb [27].

- **Similar to NPGA and NPGA II**: NSGA differs only with the ranking method used in tournament selection, just as this is the only significant way NPGA II differs from NPGA.

- **Non-dominated sorting**: in a non-dominated sort of the population, the current non-dominated subset of the population is assigned rank 0 and is then temporarily removed from consideration. The remaining population is then evaluated to determine another non-dominated subset, which is then given rank 1 and removed from consideration. This process continues until the entire population has been ranked [13].
• **Outdated:** like NPGA, which was developed around the same time, NSGA includes some fundamental MOEA components, but is now surpassed by other state-of-the-art algorithms.

6 **NSGA II (2000)**

• **Introduced** by Deb, Pratap, Agarwal, and Meyarivan [10].

• **Significantly different** from the original NSGA. From [9], “In most aspects, this algorithm does not have much similarity with the original NSGA, but the authors kept the name NSGA-II to highlight its genesis and place of origin.”

• NSGA II addresses the following all of the following drawbacks of NSGA

  – *High computational complexity of non-dominated sorting:* The non-dominated sorting algorithm from NSGA is $O(MN^3)$, where $M$ is the number of objectives and $N$ is the population size.

  – *Lack of elitism:* Results show that elitism can speed up the performance of the GA significantly and can help prevent the loss of good solutions [20, 23, 25, 28].

  – *Need for specifying the sharing parameter $\sigma_{\text{share}}$:*) The main problem with fitness sharing is that it requires the specification of the sharing parameter, which can significantly effect performance. A parameterless diversity mechanism solves this problem.

• **NSGA II is currently one of the most popular MOEAs**

• **Noisy Environments:** a head-to-head comparison between NSGA II and SPEA II [2] in noisy environments found the following: “The results show that SPEA2 outperforms NSGA2 in the early generations. NSGA2, however, is superior during latter generations regardless of the level of noise presence in the problem.” Also, there is a modified version of NSGA II designed specifically for noisy environments [1].

7 **SPEA - Strength Pareto Evolutionary Algorithm (1998)**

• **Introduced** by Zitzler and Thiele [30]
SPEA is similar to MOEAs that came before it (such as NPGA and NSGA) in the following ways:

- stores the Pareto-optimal solutions found so far externally (archiving)
- uses Pareto dominance to assign fitness values to individuals
- performs clustering to reduce the number of nondominated solutions stored without destroying the characteristics of the Pareto-optimal front

Unique aspects of SPEA at the time:

- The fitness of an individual is determined from the only the archive of nondominated solutions; whether members of the population dominate each other is irrelevant
- All solutions in the archive participate in selection
- A new niching method is used that does not require any fitness sharing parameters

SPEA has been a very successful algorithms and has resulted in numerous improved algorithms, such as SPEA2, SPEA2+, and ISPEA (all of which are listed below.)

8 SPEA2 (2001)

- Introduced by Zitzler, Laumanns, and Thiele [29]

Improvements of SPEA2 over SPEA:

- An improved fitness assignment scheme is used, which takes into account, for each individual, how many individuals it dominates and it is dominated by
- A nearest neighbor density estimation technique is incorporated which allows a more precise guidance of the search process (increased diversity)
- A new archive truncation method guarantees the preservation of boundary solutions
• SPEA2 differs from SPEA, NSGA-II, and PESA only in how it does fitness assignment and selection (this is a common theme among the most popular MOEAs.)

• Performance comparisons from the original paper:

  – SPEA2 is much better than the original SPEA in all respects on all test problems
  – SPEA2 is better than PESA in both the diversity and final quality of solutions
  – PESA converged faster on some test problem, most likely due to a higher elitism intensity
  – Comparing SPEA2 and NSGA-II: SPEA2 had a “better distribution”, i.e. less clustering, but NSGA-II had a “broader range” of solutions, i.e. found solutions closer to the outlying edges of the Pareto-optimal front
  – SPEA2 seems to outperform NSGA-II and PESA in high-dimensional objective spaces

9 **SPEA2+ (2004)**

• **Introduced** by Kim, Hiroyasu, Miki, and Watanabe [16]

• Improves on SPEA2 by adding two things:

  – A more effective crossover mechanism
  – An algorithm to maintain diversity in both the objective and decision variable spaces (previous algorithms have focused only on diversity in the objective space)

• Specifically, this is achieved with the following operations:

  – *Neighborhood Crossover*: crosses over individuals close to each other in objective space (may help maintain building blocks?)
  – *Strong elitism*: all archived individuals (non-dominated solutions) participate in the selection operators (usually tournament selection)
  – *Maintaining two archives*: one to maintain diversity in the objective space, the other in the decision space
• Performance claims:
  – When compared to SPEA2, SPEA2+ was shown to come closer to the true Pareto-optimal front on some problems.
  – With respect to the objective space, SPEA2+ was shown to have a more diverse set of solutions than NSGA-II, attributed to the neighborhood crossover operator. However, NSGA-II seems to have had a more uniform distribution of solutions. This is opposite of what was observed from the original comparison of SPEA2 to NSGA-II, indicating that there may be a tradeoff relationship between the diversity and uniformity of the solutions (?).
  – As expected, SPEA2+ showed a greater diversity in decision variable space than SPEA2 (though this was a “relatively simple” problem with a three-dimensional decision space).

10 ISPEA - Immunity SPEA (2003)
• Introduced by Hongyun and Sanyang [14]
• The basic idea seems to be: (1) identify good building blocks (combinations of parts of the genomic representation); (2) “Vaccinate” some of the population with these identified building blocks.
• Good performance is presented in the paper, but no details on how to actually perform these operations is given. Unless those details can be found, this algorithm is unrealizable.

11 PAES - Pareto Archived Evolution Strategy (2000)
• Introduced by Knowles and Corne [19]
• Is argued that PAES represents the “simplest possible nontrivial algorithm capable of generating diverse solutions in the Pareto-optimal set.”
• Has three forms, (1+1)-PAES, (1+λ)-PAES, and (µ + λ)-PAES. The notation refers to: (population size + number of new solutions per generation)
• (1+1)-PAES and (1+λ)-PAES perform only local search, i.e. they keep only one “current” solution, instead of a population of searches (basically performing multiobjective hill climbing). The (μ+λ) version maintains a population.

• PAES introduces a new crowding procedure, the *adaptive grid algorithm*. It is superior to previous niching methods in two ways: (1) Its computational cost is lower; (2) It is adaptive and does not require the critical setting of a niche-size parameter.

• The comparable performance of this simple algorithm to other more complex ones is attributed to the use of an archive of nondominated solutions, which was uncommon at the time but has been incorporated into most state-of-the-art algorithms.

• PAES is good on landscapes that are “well behaved”, i.e. not multimodal or deceptive.


• **Introduced** by Knowles and Corne [18]

• Combines the local search strategies used by PAES with population and recombination strategies.

• The population based strategies allow M-PAES to outperform PAES on multi-modal and deceptive landscapes.

• Mimetic algorithms are also known as genetic local search, hybrid genetic algorithms, and cultural algorithms.

• Much of the success of mimetic algorithms relies on the global convexity of the search space.

• The local and global search phases of M-PAES are partially independent, and each maintains their own archive of non-dominated solutions.

• M-PAES is highly elitist in both its local and recombination search methods.

• Performance Results: M-PAES compared with (1+1)-PAES, SPEA, and a single-objective optimizer (only well behaved, convex functions are tested).
– M-PAES outperforms (1+1)-PAES in all regards
– Outperforms SPEA as presented in [31]
– Comparable performance with SPEA as modified by [18] (using new parameters found to be superior on the tested functions)
– M-PAES found better solutions in the “middle” of the objective space and the single objective function found solutions farther out on the edges of the Pareto front (as expected)

13 PESA - Pareto Envelope-based Selection Algorithm (2000)

• Introduced by Corne, Knowles and Oates [8]
• The main attraction of PESA is the integration of selection and diversity management into one technique (selection and archiving are based on the diversity maintaining crowding measure)
• PESA incorporates ideas from SPEA and PAES
• Performance Results: PESA was compared to PAES and SPEA. Six different test functions were used and the best algorithm for each function was identified
  – PESA was the best on 3 functions and tied for best with SPEA on 2
  – SPEA was the best on 1 function and tied for best with PESA on 2
  – PAES was clearly the worst of the tested algorithms

14 PESA II (2001)

• Introduced by Corne, Jerram, Knowles and Oates [7]
• The main improvement of PESA is the use of region based selection. Instead of selecting individuals, hyperboxes in the objective space are selected. The hyperboxes represent a set of possible solutions. Selecting a hyperbox results in the selection of a random individual from that hyperbox. This was shown to “ensure a better spread of development along the Pareto frontier than individual based selection.”
The original source for a good argument of why region based selection is better than individual based selection.

The basic argument is as follows: When there is a non-uniform distribution of individuals, containing large clusters, tournament selection is more likely to choose individuals from those large clusters (because there are so many of them). Region based selection, however, will treat the large clusters and the isolated individuals as equal groups, with equal probability of participating in each tournament.

- **A comparison of methods to determine an individuals degree of separation:** comparing methods from PAES, PESA, NSGA-II, and SPEA
  - PAES and PESA use hyperbox counts (number of individuals in a region)
  - NSGA-II uses the product of the distances to the nearest neighbors along each objective dimension
  - SPEA uses the number of other individuals dominated

- All of the above methods are *individual based*, in that they select a particular individual. The differences are in the calculation of the selection probability for each individual.

- Performance results: PESA II was compared to PAES, SPEA, and PESA. It was shown to be superior to all of them. No comparison was done with NSGA-II in the original paper (perhaps other literature provides such a comparison).

- PESA II requires the specification of the hyperbox dimensions. This causes problems similar to those introduced when specifying the niching parameters of other algorithms. Results may be very sensitive to this parameter.


- **Introduced** by Coello and Pulido [5]

- Uses a very small population (4 individuals) and a re-initialization process
• Aims to be a very fast converging algorithm with low computational cost

• Argued to be the first micro-GA for multiobjective optimization, though population based forms of PAES could be considered a micro-GA (but PAES is geared more toward including local search than most GAs).

• Authors list three types of elitism that they use, each of which has been shown in some form in other algorithms

  1. Storing nondominated solutions in an external archive (common to most successful algorithms)
  2. Carrying over the best individuals between re-initializations of the small population
  3. Allowing individuals from the nondominated archive to participate in the active search (also common in successful algorithms)

• No performance comparisons were made in the original source (though see below for performance comparisons made between micro-GA2 and other MOEAs)

16 Micro-GA2 (2003) a.k.a. $\mu$GA$^2$

• Introduced by Pulido and Coello [24]

• Intended to be a step toward online adaptation, i.e. self-adapting parameters. The setting of sensitive algorithmic parameters has always been a major concern for evolutionary algorithms and MOEA usually have more parameters than traditional GAs (e.g. niche radius, fitness sharing threshold, etc.)

• Has two main stages: exploration and exploitation. The main difference between the two stages is in the relative weighting of mutation and crossover. In the exploration stage, mutation is dominates; in the exploitation stage, crossover dominates.

• The parameters that are adaptive are: crossover rate, size of the population, the percentage of the population that is “replaceable” (used to control the amount of elitism), rate at which the replaceable portion of the population is replaced, total number of iterations, number of subdivisions of the grid (hyperboxes used to control clustering.)
• Some other parameters, such as which crossover method to use, are decided online by testing them in a set of parallel micro-populations.

• Performance comparison: micro-GA2 was compared to micro-GA, NSGA-II, and PAES (compared using generational distance, error ratio, and spacing)
  – micro-GA2 consistently outperformed both micro-GA and PAES
  – NSGA-II outperformed micro-GA2 on some test functions and was nearly even on all others (NSGA-II may still be superior)


• Introduced by Cochran, Horng, and Fowler [3]

• Uses a serial two stage process, each of which is based on an older (now outdated) multi-objective algorithm:
  – Stage 1: based on MOGA [21]. Uses a weight vector to unify the multiple objectives into a single objective. Uses a new, random weight vector for each generation to promote a distribution of solutions on the Pareto front, instead of a single solution to a specific weight of objectives.
  – Stage 2: based on VEGA [26]. Uses $n$ parallel populations, each one optimizing one of the $n$ objectives. Promotes finding the extreme solutions of the Pareto front.

• No direct comparison of MPGA to current state of the art MOEAs (such as NSGA-II and SPEA 2) was found. Because it is based on two algorithms that have been shown to be inferior to the best current algorithms, MPGA is likely not the best solution on its own. However, the idea of a multi-stage search, each stage with its own strengths, is a valuable one.

• When using a multi-stage approach, an important concern is the choice of the turning point, i.e. when to switch to the next stage. In this case, the proper turning point was determined experimentally and was found to be around one-half of the total fitness evaluations (i.e. each stage ran for about half time).

- **Introduced** by Coello and Becerra [4]
  - The first algorithm of its kind: a cultural algorithm for multiobjective optimization adopting Pareto ranking and elitism
  - Cultural algorithms split the inheritance process into two levels: the micro-evolutionary level, and the macro-evolutionary level.
  - Cultural algorithms operate in two spaces, the normal population space, and an additional belief space. Knowledge gained by the evolution in the population space is stored in the belief space and is used to influence further evolution.
  - The belief space in this case contains two parts:
    1. The phenotypic normative part: the lower and upper bounds of the intervals for each objective function
    2. A grid to prevent solution clustering (a variation on the grid proposed by Knowles and Corne [19])
  - The belief space therefor represents and adaptive objective space grid (the adaptive nature of the grid is the primary difference of this algorithm to previous MOEAs)
  - Performance comparison: CAEP was compared to NSGA-II. They report “competitive behavior” between the two, indicating that CAEP was not found to be better than NSGA-II in any significant way.

19 MOPSO - Multi-Objective Particle Swarm Optimization (2003)

- **Introduced** by Coello, Pulido, and Lechuga [6]
  - An extension of previous Particle Swarm Optimization (PSO) algorithms from single to multi-objective
  - Based on flocking behavior (of birds, fish, etc.) where the direction of movement of an individual is effected by the locations and movement of neighboring individuals
• PSO has been shown to provide very short convergence times. The main advantages of MOPSO are its short convergence times and low computational overhead.

• MOPSO uses a mutation operator that acts on both the individuals and on the ranges of each decision variable to be explored. This unique method attempts to fully search the decision space, not simply maintain diversity in the objective space.

• Extensive experimental results are given, both to define the sensitivity of the parameters of the algorithm and to compare MOPSOs performance to other MOEAs.

• MOPSO uses a form of region based selection, instead of individual selection (similar to that of [7]).

• Performance comparison of MOPSO with NSGA-II, micro-GA, and PAES
  – MOPSO was the only algorithm in the study “that was able to cover the full Pareto front of all the functions used”
  – MOPSO may need the addition of a crowding operator to improve the uniformity of the solution distribution
  – Found to be “competitive” with other algorithms

• The authors provided an extensive analysis of parameters and give definite recommendations for parameter settings.

20 ParEGO - Pareto Efficient Global Optimization (2005)

• Introduced by Knowles [17]

• This is the probably the most unique MOEA out of all of the ones review here.

• A hybrid approach: uses both an evolutionary search and a detailed internal model. The model is used to select areas of the space to search that will both provide better solutions and improve the internal model (i.e. reduce the uncertainty about the search landscape)
• Designed specifically to work with very few fitness evaluations (on the order of 100 to 300)

• Appropriate for search problems with the following characteristics
  – Very expensive fitness function evaluation
  – No parallel fitness evaluation is possible
  – Noise is low
  – The search landscape is smooth (but can be multi-modal)
  – The dimensionality of the space is low-to-medium

• Converts the multiobjective search into a single objective search, using different objective weightings on successive iterations

• Shows very good performance when compared to NSGA-II for very few fitness evaluations (100 and 250)

References


