Improvement of NSGAⅢ Algorithm Based on Information Entropy

Guangxin Fan*, Xiaojie Wang

Southwest Petroleum University, Chengdu 610500, China.

*van.sing@foxmail.com

Abstract

Group search strategy and information exchange among individuals are two advantages of evolutionary algorithm in solving multi-objective optimization problems. At present, the multi-objective evolutionary algorithm based on the concept of Pareto optimization has become the main research direction of multi-objective optimization problems. This paper briefly introduces the NSGAIII algorithm for high-dimensional and multi-objective optimization in recent years, and proposes an improved nsga3 algorithm based on information entropy. It is proved to be effective and feasible by classical test problems.

Keywords

Multi-objective optimization, NSGAIII, Information entropy.

1. Introduction

In the study of multi-objective optimization, as the target dimension increases, the optimization difficulty increases exponentially. Usually we call the optimization problem of 4 or more objectives as high-dimensional multi-objective optimization. Compared with the classic multi-objective optimization, the evolution of high-dimensional multi-objectives has the following difficulties:1. Degradation of search ability Because of the increase of the target dimension, the number of non-dominated individuals in the population increases exponentially, which reduces the selection pressure during its evolution.2. The number of non-dominated solutions used to cover the entire Pareto front increases exponentially.3. It is difficult to visualize the optimal solution set.4. It is difficult to evaluate the distribution of the solution set.5. The efficiency of reorganization operation is reduced. In a larger high-dimensional space, the recombination of two distant husbands may produce offspring farther from the parent, which weakens the local search ability of the population.

2. Introduction of NSGAⅢ algorithm

In recent years, scholars have proposed many effective methods for solving high-dimensional multi-objective optimization. The well-known classic NSGAII algorithm can handle multi-objective optimization problems well, but can only handle multi-objective problems with lower dimensionality. In higher dimensions, the non-dominated individuals in the population increase exponentially, making it difficult to distinguish their individual pros and cons through the Pareto dominance relationship. On this basis, the NSGAIII algorithm is proposed.

The overall framework of NSGAIII is similar to that of NSGAII. The process is shown in the following figure.
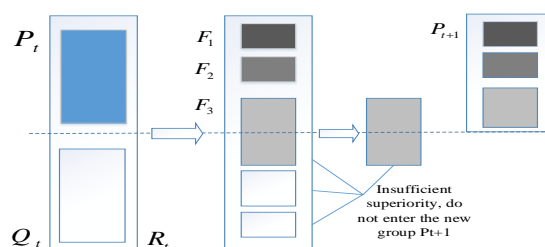

Figure 1. New population generation process

In the selection process, the non-dominated tiers are divided into different non-dominated tiers by means of non-dominated sorting. The non-dominated tiers with superiority and high priority are retained for the next generation. If it is not in the dominant layer, then it is called the critical layer. In the figure, F3 is the critical layer, and then the critical layer scholar method is used to select individuals to enter the new generation of population, so that the population size is N. The difference between NSGAIII and NSGAII is that it uses the method of reference points to select individuals, so that individuals have a better distribution.

3. NSGAⅢ algorithm based on information entropy

Information entropy is used to describe the constant degree of information and represents the probability of a certain specific information. It is a concept in information theory. In recent years, it has also been used in various algorithms or other practical applications. Professor Zheng Jinhua proposed the entropy-based MOEA (EMOEA). Inspired by this, the information entropy mechanism in the population can be used to change the crossover probability of mutation in NSGAIII.

The information entropy of the population is usually defined as the following formula:

$$E = -\sum_{i=1}^{n} q_i \log q_i \tag{1}$$

Where: $q_i = |P_i|/N$, N is the number of individuals in the population; $|P_i|$ represents the number of P_i all individuals. n is the number of individuals whose population is in the same range on the optimization objective. When $n=1$, the population information entropy was the smallest, and its value was 0;When $n=N$, the value was the largest.and its value was $\log N$.

At this time, according to the population information entropy, the crossover and mutation probability of population evolution can be rewritten as the following formula:

$$\begin{cases} P_c^1 = c_1 - \dfrac{c_1 n}{N} \\ P_m^1 = c_2 - \dfrac{c_2 n}{N} \end{cases} \tag{2}$$

Among them: c_1、c_2 are constants greater than 0 and less than 1, respectively.

In the initial stage of population evolution, the similarity of population individuals is low, increasing the crossover probability can accelerate the evolution process of the population, reduce the mutation probability of the population, and reduce the tedious calculation process; in the middle stage of the evolution process, appropriately increase the mutation probability , It can effectively prevent the algorithm iteration from falling into the local optimum; at the end of the evolution process, the individual similarity is high, in order to make the algorithm gradually converge, appropriately reduce the crossover probability and mutation probability. According to the above principles, the crossover probability is approximated as a decreasing function model in evolutionary algebra, and the mutation probability is approximated as a Cauchy distribution model:

$$\begin{cases} P_c^2 = c_3 \cos \dfrac{\pi}{2T} t \\ P_m^2 = \dfrac{1}{c_4 \pi \left[1 + (\dfrac{t - T/2}{c_4})^2 \right]} \end{cases} \tag{3}$$

Among them: t is the current evolutionary algebra; T is the total algebra of evolution; c_3 is a constant greater than 0 and less than 1; c_4 is a scale parameter, and is a constant greater than 0.

In summary, the final crossover probability and mutation probability of the improved algorithm are as follows:

$$\begin{cases} P_c^2 = c_1 - \dfrac{c_1 n}{N} c_3 \cos \dfrac{\pi}{2T} t \\ P_m^2 = \dfrac{1}{c_4 \pi \left[1 + (\dfrac{t - T/2}{c_4})^2 \right]} \end{cases} \tag{4}$$

Information entropy reflects the uniformity of individual distribution in the solution space. Using information entropy to improve the crossover probability and mutation probability of each generation of the population can improve the convergence accuracy and computing efficiency of the algorithm.

The NSGAIII algorithm based on information entropy(E-NSGAIII) flow is shown in the figure:
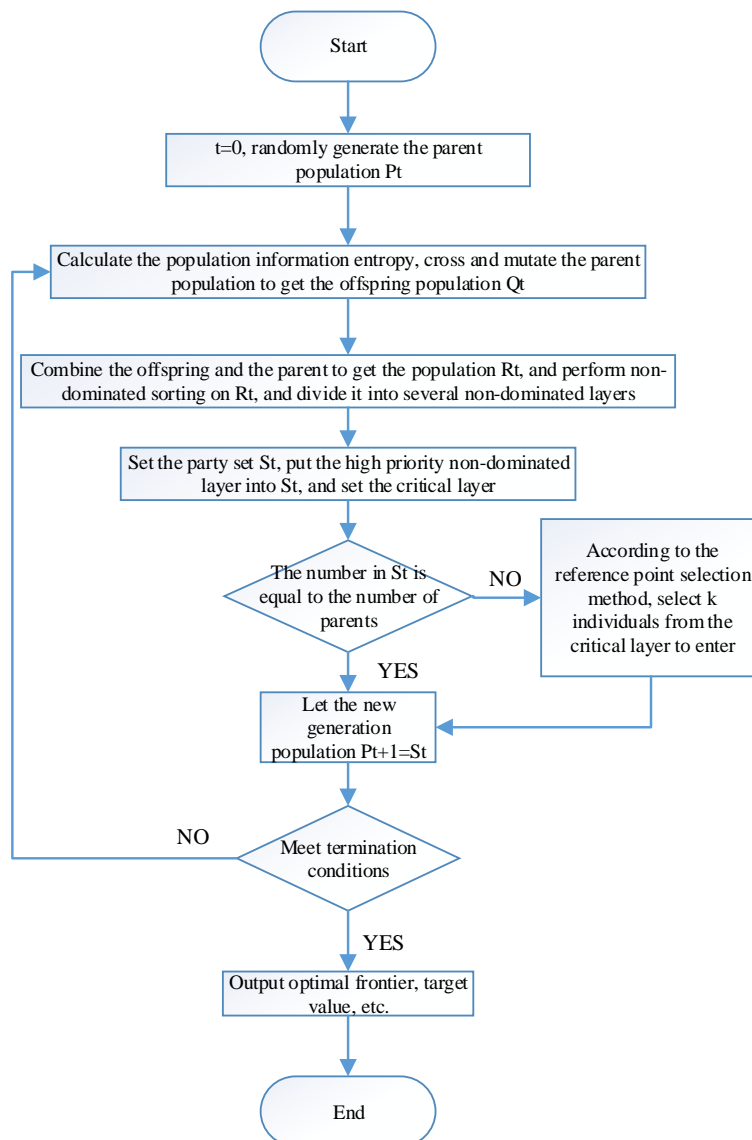
Figure 2. E-NSGAIII algorithm flow chart

4. Analysis of model parameters and results

This article chooses the test function as DTLZ1~DTLZ3 and WFG4~WFG6, and the results are shown in the following chart:
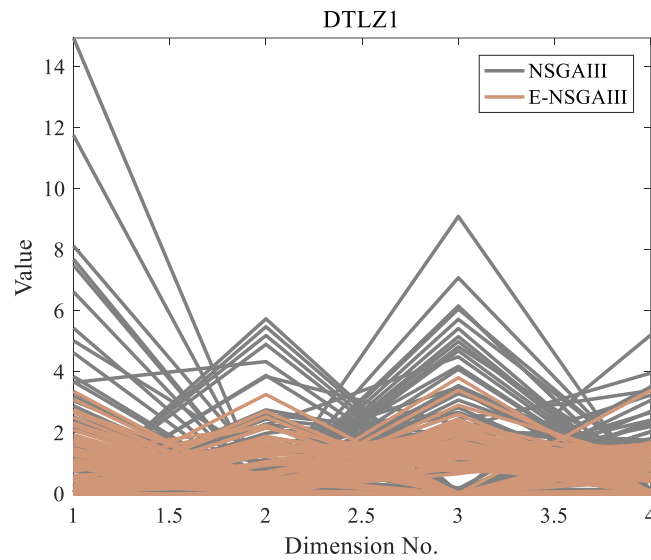
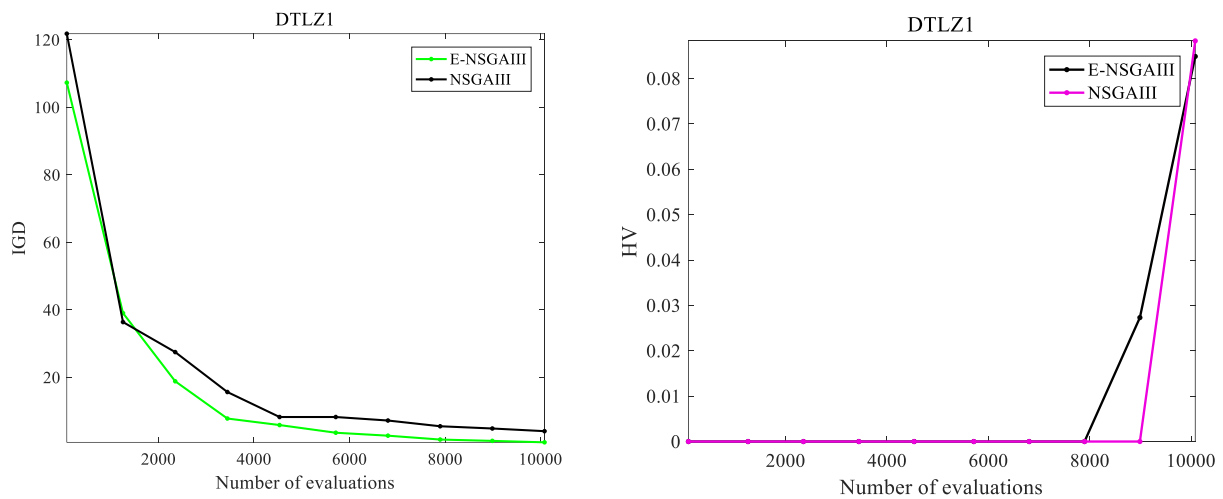Figure 3. E-NSGAIII and NSGAIII at the frontier of 4-dimensional DTLZ1

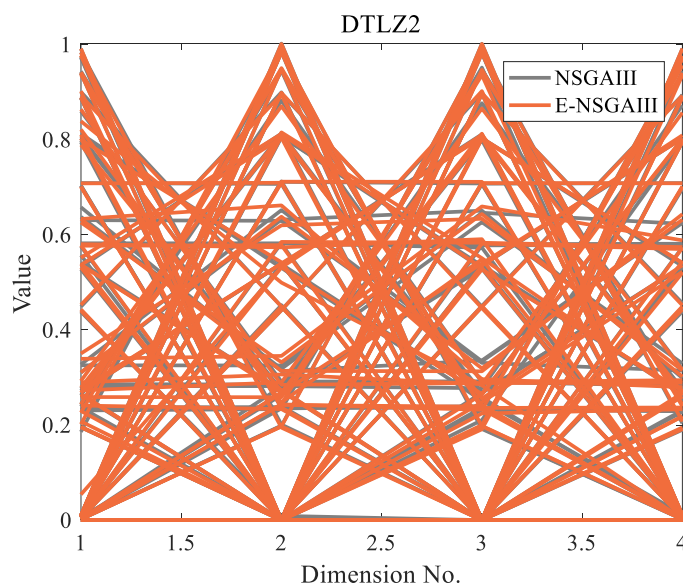Figure 4. The algorithm's IGD and HV indicators on 4-dimensional DTLZ1

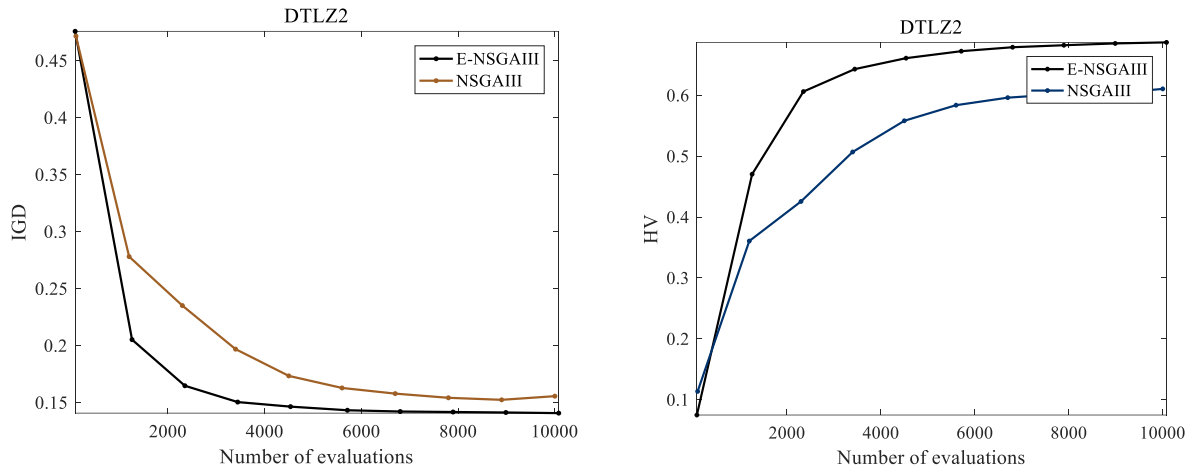Figure 5. E-NSGAIII and NSGAIII at the frontier of 4-dimensional DTLZ2

57

Figure 6.The algorithm's IGD and HV indicators on 4-dimensional DTLZ2

The improvement of NSGAIII based on information entropy is mainly to improve its computing efficiency, speed up the genetic process, and improve its convergence accuracy. For this reason, the evaluation index selects its inverse generation distance (IGD) as the comprehensive index of its convergence accuracy, and uses the running time (TM/s) as its calculation efficiency index. Choose different test questions to verify. Select the test functions DTLZ1~DTLZ3 and WFG4~WFG6, the target number is 4, the initial total group is 100, the decision variable is 20, and the evolution algebra is 1000. The test results are as follows:

Table 1. Algorithm test results comparison

Test function	Number of goals	Evaluation index	NSGAIII	E-NSGAIII
DTLZ1	4	IGD	2.9768e+01	1.4735e+01
		TM/s	1.4161e+00	9.8834e-01
DTLZ2	4	IGD	5.8057e-01	2.9660e-01
		TM/s	1.5868e+00	1.1863e+00
DTLZ3	4	IGD	9.1863e+01	7.8094e+01
		TM/s	1.3234e+00	9.7775e-01
WFG4	4	IGD	1.9582e+00	1.9162e+00
		TM/s	1.1685e+00	9.5463e-01
WFG5	4	IGD	1.9397e+00	1.9145e+00
		TM/s	1.0714e+00	1.0112e+00
WFG6	4	IGD	1.9873e+00	1.9124e+00
		TM/s	1.2735e+00	1.0413e+00

It can be seen from chart that the E-NSGAIII algorithm has improved convergence accuracy and computational efficiency, and its solution set can better approximate the real Pareto frontier compared with the NSGAIII.

5. Conclusion

This paper uses the related theory of information entropy to improve the crossover and mutation probability in the NSGA3 algorithm, and verifies it through the test function, which proves the effectiveness, feasibility and superiority of the improved algorithm Ensga. Through the test results, it is found that Ensga algorithm has better solving efficiency and better performance indicators than NSGA3 algorithm.

References

[1] Deb K. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-2[J]. IEEE Trans. Evol. Comput, 2002, 6:849--858.

[2] Seada H, Deb K. Effect of selection operator on NSGA-III in single, multi, and many-objective optimization[C]// Evolutionary Computation. IEEE, 2015.

[3] Deb, K. Multi-objective Genetic Algorithms: Problem Difficulties and Construction of Test Problems [J]. Evolutionary computation, 1999, 7(3): 205-230

[4] Deb K. Multi-objective optimization using evolutionary algorithms[M]. John Wiley & Sons, Inc, 2001.

[5] Huband S, Hingston P, Barone L, et al. A review of multi-objective test problems and a scalable problem toolkit[J]. IEEE Transactions on Evolutionary Computation, 2006, 10(5):477-506.

[6] Deb K, Thiele L, Laumanns M, et al. Scalable multi-objective optimization test problem[C]// Evolutionary Computation, 2002:825-830.

[7] Deb K, Thiele L, Laumanns M, et al. Scalable Test Problem for Evolutionary Multiobjective Optimization [M]. Springer London, 2005.

[8] Fu Zuyun. Information theory: basic theory and application [M]. Electronic Industry Press, 2011