

Improving evolutionary algorithms with information feedback model for large-scale many-objective optimization

Yong Wang1 · Qian Zhang¹ · Gai-Ge Wang¹

Accepted: 4 July 2022 / Published online: 6 September 2022 © The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Recently, many evolutionary algorithms have been proposed. Compared to other algorithms, the core of the many-objective evolutionary algorithm using a one-by-one selection strategy is to select offspring one by one in environmental selection. However, it does not perform well in resolving large-scale many-objective optimization problems. In addition, a large amount of meaningful information in the population of the previous iteration is not retained. The information feedback model is an effective strategy to reuse the information from previous populations and integrate it into the update process of the offspring. Based on the original algorithm, this paper proposes a series of many-objective evolutionary algorithms, including six new algorithms. Experiments were carried out in three different aspects. Using the same nine benchmark problems, we compared the original algorithm with six new algorithms. Algorithms with excellent performance were selected and compared with the latest studies using the information feedback model from two aspects. Then, the best one was selected for comparison with six state-of-the-art many-objective evolutionary algorithms. Additionally, non-parametric statistical tests were conducted to evaluate the different algorithms. The comparison, with up to 15 objectives and 1500 decision variables, showed that the proposed algorithm achieved the best performance, indicating its strong competitiveness.

Keywords Many-objective · Information feedback model · One-by-one selection · Evolutionary algorithms · Large-scale optimization

1 Introduction

Owing to the needs of engineering applications and scientific experimental research, there are many multi-objective optimization problems (MOPs) in real life. The main optimization goal is to maximize or minimize the final value with as few resources as possible [\[1\]](#page-32-0). The number of objective functions of MOPs increases to two or three, unlike in the single-objective optimization problem [\[2\]](#page-32-1). However, the objective function for many optimization problems is

- Gai-Ge Wang [wgg@ouc.edu.cn](mailto: wgg@ouc.edu.cn)

> Yong Wang [markwy@126.com](mailto: markwy@126.com)

Qian Zhang [zhangqaq@stu.ouc.edu.cn](mailto: zhangqaq@stu.ouc.edu.cn)

¹ School of Computer Science and Technology, Ocean University of China, Qingdao, 266100, China

not limited to three or fewer. As a result, many-objective optimization problems (MaOPs) with more than three goals arise [\[3\]](#page-32-2). There is often no ideal solution to achieve optimal values for all goals. At present, methods to solve optimization problems are generally divided into two categories: traditional mathematical optimization methods (e.g., gradient descent and Newton) and evolutionary algorithms [\[4\]](#page-32-3).

Traditional optimization algorithms that use information on the first derivative and second derivatives of related functions are often unsuitable for practical application problems and cannot guarantee that the final approximate optimal solution set can be evenly distributed and converged [\[5\]](#page-32-4). To solve this problem, many researchers have proposed many evolutionary algorithms. These algorithms draw inspiration from biological evolution and guide the process of finding solutions by simulating the process of biological evolution, using the survival of the fittest principle in genetics [\[6,](#page-32-5) [7\]](#page-32-6). Such as genetic algorithm (GA) [\[8\]](#page-32-7), particle swarm optimization (PSO) [\[9\]](#page-32-8), artificial bee colony (ABC) [\[10\]](#page-32-9), ant colony algorithm (ACO) [\[11\]](#page-32-10) and differential evolution (DE) [\[12\]](#page-32-11).

Many scholars have proposed numerous multi-objective optimization algorithms (MOEAs) to solve MOPs. The most famous ones are the decomposition-based multiobjective evolutionary algorithms (MOEA/D) [\[13\]](#page-32-12), genetic algorithms based on non-dominated sorting and elite selection (NSGA-II) [\[14\]](#page-32-13), evolutionary algorithms using neighborhood environment selection (SPEA2) [\[15\]](#page-32-14), etc. With the advent of 5G era and the rise of deep learning, evolutionary algorithms have applications in many emerging fields [\[16\]](#page-32-15). They are currently widely used in economic dispatching [\[17\]](#page-32-16), code detection [\[18\]](#page-33-0), engineering optimization [\[19\]](#page-33-1), face reconstruction [\[20\]](#page-33-2), jobshop scheduling $[21-23]$ $[21-23]$, high performance computing $[24]$, computer network [\[25\]](#page-33-6), and other fields. These MOEAs have been proved successful in solving most MOPs. However, the non-dominated solution set obtained by these algorithms often cannot effectively approximate the real Pareto front of MaOPs with optimization goals greater than three $[26]$. Therefore, considering the improvement of the convergence and diversity of the algorithms, this paper examines the many-objective evolutionary algorithm using a one-by-one selection (1by1EA) [\[27\]](#page-33-8), aiming to improve its optimization performance in large-scale MaOPs. The method uses a convergence index to select the solutions in the current population one by one, adopts a distribution index based on cosine similarity to evaluate the similarity between solutions, and proposes the boundary maintenance mechanism of corner solutions to maintain diversity.

However, few researchers have focused on the utility of prior knowledge and historical information. Many have adopted the method of retaining the best individual in the process of each iteration and discarding all other individuals. This is not conducive to the development of population diversity and the preservation of information. More importantly, most MOEAs focus on solving optimization problems with less than 100 decision variables. The problem of more than a thousand decision variables in engineering application research may also exist. Therefore, we have to enhance 1by1EA to improve the result. Therefore, this paper introduces the information feedback model (IFM) [\[28\]](#page-33-9) into the framework of 1by1EA and proposes a new optimization algorithm called 1by1EA-IFM, which solves large-scale MaOPs. Unlike most existing algorithms, this algorithm retains the historical information of individuals in the population before selecting offspring individuals. The experiment was carried out in three aspects. First, the performance of the improved six algorithms was compared with the original algorithm. Second, algorithms with excellent performance were selected and compared with current IFM research from two levels. Finally, we chose the one with the best performance and compared it with the other six state-of-the-art algorithms from a longitudinal perspective.

The main contributions of this work can be summarized as follows:

- (1) A novel framework, called 1by1EA-IFM is proposed, where individuals from historical populations are retained in two ways: either in a fixed location or randomly retained. The individual performance is determined by the fitness function value. In each retaining approach, *k* individuals are selected $(k \in \{1,2,3\})$. Based on the value of *k*, IFM includes six models (i.e., M-F1, M-F2, M-F3, M-R1, M-R2, and M-R3), which are combined with 1by1EA to present six algorithms, namely 1by1EA-F1, 1by1EA-F2, 1by1EA-F3, 1by1EA-R1, 1by1EA-R2, and 1by1EA-R3. An advantage of IFM is that it can efficiently enhance the diversity of the population by reusing some information from historical populations.
- (2) Experiments were performed on large-scale MaOPs, where the number of decision variables reaches up to 1500. By controlling the number of decision variables, we can improve the scale of the problem, so as to test whether the algorithms can effectively target large-scale problems. In all the validation experiments, the comparison of the indicator values involving the different algorithms was carried out using the Mann-Whitney-Wilcoxon rank-sum test, which can determine whether one algorithm has statistical difference with other algorithms. In addition, with the intention of obtaining rigorous conclusions, two different statistical tests, Friedman's non-parametric test and Holm's post-hoc test, were conducted.

The rest of this paper is organized as follows. In Section [2,](#page-1-0) we introduce the work related to this research. Section [3](#page-3-0) provides some basic concepts about multi/many-objective optimization, as well as the original 1by1EA and IFM. The proposed 1by1EA-IFM is described in detail in Section [4.](#page-6-0) The settings and results related to the experiment are described in Section [5.](#page-7-0) Section [6](#page-32-17) provides a summary of this paper.

2 Related work

There are a variety of optimization problems with more than three objectives in social life, and the corresponding solution is the many-objective evolutionary algorithms (MaOEAs). For most evolutionary algorithms that rely on Pareto dominance, the non-dominant solution sets often cannot effectively approximate the real Pareto front of MaOPs. Therefore, redefining the dominance relationship is an effective measure to solve MaOPs for MOEAs. The main studies were ε domination [\[29\]](#page-33-10), fuzzy Pareto domination [\[30\]](#page-33-11) and so on. Chhabra et al. [\[31\]](#page-33-12) proposed a fuzzy Paretodriven artificial swarm algorithm (FP-ABC) to better solve for MaOPs. In the FP-ABC algorithm, two external files are combined into the artificial bee colony algorithm to improve the performance of the ABC algorithm. In 2022, Wu et al. [\[29\]](#page-33-10) introduced a new *ε* dominance relation and proposed a new MaOEAs (*ε*-Two Arch2) to update the individuals in diversity archive. The common problem with these algorithms is that they may cause the population to converge to a sub-region of the Pareto optimal front.

Unlike the above algorithms based on improving dominance relationship, decomposition-based MaOEAs decompose MaOPs into several sub-problems for simultaneous optimization, and the goal of the problem is aggregated into different scalar functions. These algorithms guide the individual to search in the direction near the Pareto front by minimizing these scalar function values. In 2018, Zheng et al. [\[32\]](#page-33-13) redesigned the weight vectors used in the sub-problem and proposed a new weighted mixture-style method to enhance MOEA/D. Lucas et al. [\[33\]](#page-33-14) introduced a new MOEA/D with uniformly randomly adaptive weights (MOEA/D-URAW) to deal with the limitation of algorithm independent of problem geometry. However, when the Pareto front of the problem is too complex and inconsistent with the distribution of the weight vector, the distribution of the population obtained by MOEAs is often poor.

Different from the above algorithms, MaOEA/IGD [\[34\]](#page-33-15) is a many-objective evolutionary algorithm based on the IGD index. This algorithm first used a singleobjective evolutionary algorithm to estimate the range of Pareto front, and then established a hyper-plane in the target space according to this range. Bader et al. [\[35\]](#page-33-16) used Monte Carlo approximation based hypervolume indicators for environmental selection, and proposed a hypervolume estimation algorithm (HypE). One of the biggest problems with the above indicator-based methods is that the computational complexity of the indicator can be too high, thus, the algorithm speed may be generally slower than that of ordinary algorithms.

In addition to MaOEAs covered in the above classifications, the many-objective evolutionary methods have shown some new characteristics in recent years. The selection phrase plays a key role in most algorithms, taking into account both the convergence and the distribution of the solution. However, the increase in the number of objectives and decision variables has brought huge challenge to MaOEAs, such as Pareto dominance ineffectiveness and the conflict problem of convergence and diversity [\[36\]](#page-33-17).

2.1 Selection strategy

Many scholars have studied MaOEAs based on selection strategies. In the first few years of evolutionary algorithms'

development, in order to avoid loss of external solutions leading to reduced diversity, Zitzler et al. [\[15\]](#page-32-14) proposed a strength Pareto evolutionary algorithm (SPEA2), which used density estimation, archive truncation, and fine-grained fitness allocation as environmental selection to solve MaOPs. Deb et al. [\[37\]](#page-33-18) used a reference point strategy-based nondominated sorting algorithm (NSGA-III) to improve the selection process of NSGA-II to choose individuals with good convergence by using a set of pre-defined references point as the standard. Recently, many novel studies have emerged one after another. In 2021, Liu et al. [\[38\]](#page-33-19) proposed a manyobjective evolutionary algorithm based on decomposition with correlative selection mechanism (MOEA/D-CSM) to find its correlative individuals for each reference point as soon as possible to maintain the diversity of the population. In the same year, Palakonda et al. [\[39\]](#page-33-20) proposed an ensemble framework (ENMOEA) in which mating and environmental selections of diverse MOEAs are combined. The framework demonstrates the scalability of the algorithm with the addition of a selection strategy. In this year, in order to solve discrete MaOPs, Zhao et al. [\[40\]](#page-33-21) adopted an adaptive selection strategy to improve the convergence performance of decomposition-based ACO by using different reference points. In the end, this algorithm can effectively improve the quality of optimization.

Although the improvement of the selection strategy makes the overall performance of the algorithm better, trying more strategies can further improve the universality of the algorithm. However, the majority of researchers have not considered the scale of the problem, but have only solved the problem of the number of variables in smalland medium-scale decisions. Therefore, in-depth study of large-scale problems is needs.

2.2 Large-scale optimization

In order to deal with the problem of a large number of decision variables in practical, many researchers have made relevant studies. Chen et al. [\[41\]](#page-33-22) introduced a scalable small subpopulations based covariance matrix adaptation evolution strategy (S3-CMA-ES). The algorithm used a series of small subpopulations to approximate the Pareto optimal solution, and introduced a variety of diversity improvement strategy to solve the MaOPs of large-scale decision variables. He et al. [\[42\]](#page-33-23) embedded adaptive offspring generation method in a MOEA framework (DGEA) and proposed a pre-selection strategy to select parents and used them to construct a direction vector in the decision space to propagate offspring. Tian et al. [\[43\]](#page-33-24) combined a competitive swarm optimizer with large-scale multi-objective optimization (LMOCSO), the proposed algorithm used a new particle update strategy based on two stages to update the particles.

The focus of previous research has been on the scale of the problem, but the operation of the specific details of the algorithm has not been studied in-depth, such as how to better coordinate the environmental selection and mating selection, or how to use historical information more effectively. However, the emergence of the above algorithms provides an idea for the development of this paper.

Considering the influence of selection strategy, this paper introduces IFM into 1by1EA. Instead of selecting the best individual each time, IFM selects individuals from the previous iteration in a fixed or random manner, and then, they are used for the generation of offspring, thus increasing the diversity of the population. Prior to this, researchers have conducted research on IFM. For example, Gu et al. [\[44\]](#page-33-25) and Zhang et al. [\[45\]](#page-33-26) used IFM to improve NSGA-III and MOEA/D, respectively. The proposed algorithms are used to solve large-scale MaOPs. Therefore, the new algorithm is compared with those in the two above mentioned studies. For large-scale research, we chose DGEA and LMOCSO which are also used to solve large-scale problems, and the remaining four state-of-the-art algorithms, which have been proposed in the last 3 years. IFM improves the competitiveness and persuasiveness of 1by1EA to solve problems with large-scale decision variables.

3 Preliminaries

3.1 Basic definitions

MOPs are a common problem in many areas of the real world. Assuming they are a minimization problem, the optimization goal is to minimize all objective functions as much as possible. MOPs can be defined as [\[46\]](#page-33-27):

$$
minF(x) = [f1(x), f2(x), ..., fm(x)]T
$$

s.t. $x \in \Omega$, $f \in Rm$ (1)

where *m* is the number of objective functions and Ω is decision space. *x* is an *n*-dimensional decision variable, that is, $x = [x_1, x_2, \dots, x_n]^T \in \Omega$, it includes possible solutions to the problem. For the set of *m*-dimensional objective functions $F : f \rightarrow R^m$, it matches the *n*-dimensional decision space and the *m*-dimensional objective space. When $m > 3$, this problem can be defined as an MaOP.

The following concepts have been widely popularized:

Pareto dominance: For any two solutions x_1 , x_2 in (1) , x_1 Pareto dominates x_2 if they satisfy the following conditions, denoted as $x_1 \prec x_2$.

$$
f_i(x_1) \le f_i(x_2), \forall i \in \{1, 2, ..., m\}
$$

$$
f_j(x_1) < f_j(x_2), \exists j \in \{1, 2, ..., m\}
$$
 (2)

- **Pareto optimal set:** For $x^* \in \Omega$ in (1), if there is no solution $x_1 \in \Omega$ satisfying $x_1 \prec x^*$, then x^* is known as the Pareto optimal solution. All of these solutions come together to form the Pareto optimal set (PS).
- **Pareto optimal front**: The set of objective value vectors corresponding to each solution in PS is called Pareto optimal front (PF).

$$
PF = \{F(x) \mid x \in PS\}
$$
\n⁽³⁾

3.2 1by1EA

In order to balance the convergence and diversity of solutions in the high-dimensional target space, 1by1EA was proposed [\[27\]](#page-33-8) and used to solve MaOPs.

The main contribution of this algorithm is that it uses a convergence index to select the solutions in the current population one by one, and proposes a distribution index instead of the Euclidean distance to evaluate the distance between the solutions in the high-dimensional space. It can be used to choose the neighboring solution of the selected solution, so as to weaken it by using niche technology. Unlike most MaOEAs, a boundary maintenance mechanism ensures that the corner solution is not easily discarded. A corner solution can be defined as the individual with the smallest scalar value aggregated by *k* objectives in the current population, and the number of *k* is less than the total number of objectives. The general operating framework of the algorithm is given in Algorithm 1. First, *N* individuals are randomly generated to form the initial parent population (*P*). Then, mating selection is performed on *P* to select parents for producing offspring, which iterates until *N* offspring are produced. The offspring form the offspring population *Q*. After the variation operations are performed on Q , Q_1 is obtained, and P and Q_1 are combined to form *K*. The convergence index and distribution index of each solution are calculated in *K*, and then, the one-by-one selection method is used to select *N* individuals from *K* to constitute the initial population of the next generation.

8: end while

When calculating the convergence index, each target in the high-dimensional space is treated equally and aggregated into a scalar. The equation is as follows [\[27\]](#page-33-8):

$$
c(x) = agg(f_1(x), f_2(x), \dots, f_m(x))
$$
\n(4)

the distribution index takes the form of a vector, where each element represents the distance between a solution and other solutions in the population. Based on the fact that Euclidean distance is unsuitable for the distribution in high-dimensional space, cosine similarity can use the cosine value of the angle between two vectors to measure the similarity. The distribution index can be expressed as follows [\[47\]](#page-33-28):

$$
d(x_i) = (d_1(x_i), \dots, d_{|k|}(x_i)), i = 1, \dots, |k|
$$
 (5)

where

$$
d_j(x_i) = 1 - \cos(\theta_{ij}), \, j = 1, \dots, |k| \tag{6}
$$

is the distance between x_i and x_j . The smaller the value, the more similar the distribution between the two solutions. This indicator can effectively remove points close to the coordinate axis instead of on the PF.

As the key operation of 1by1EA, the execution process of the one-by-one selection strategy in environment selection can be described as follows:

Step 1 Boundary maintenance. Corner individuals are selected from K in set K_s .

Step 2 Determine the set K_s .

- Step 2.1 The convergence index is calculated for the remaining individuals in K and they are put into K_s one by one according to the one-by-one selection strategy.
- Step 2.2 The distribution index is used to measure which solutions are close to the individuals in K_s , and then, the solutions and those dominated by these solutions are de-emphasized. Thus, the non-dominated solutions are retained.
- Step 2.3 Stopping criterion. If *K* is not null, repeat Step 2.1 and Step 2.2, and if *K* is null, go to Step 2.4.
- Step 2.4 Selection. If the number of solutions in K_s is greater than *N*, then the first *N* individuals are the initial parents of the next generation. If the number of solutions in K_s is less than N , the de-emphasized and dominated solutions will be selected and filled into K_s according to the principle of survival of the fittest until the number is *N*.

Step 3 $P = K_s$. K_s is the next generation of parents.

Before calculating the two index values, the normalization of each individual in *K* ensures that only corner solutions are retained in *K*. Previous studies have shown that the normalization operation can effectively deal with the problem of dimensionality curse [\[37\]](#page-33-18). For other details about the algorithm, please refer to the original paper [\[27\]](#page-33-8).

3.3 Information feedback model

As mentioned before, IFM can use a simple fitness weighting method to extract and fully use the information in the previous iteration. In the previous iteration, a fixed position or random method can be used to select *k* individuals from the population. The method of selecting individuals in a fixed manner is classified as F, and selecting individuals in a random manner is classified as R. If *k* is 1, the model can be defined as M-F1, and six IFMs can be obtained.

In M-F, individuals are selected in a fixed manner, which means that for individuals who need to be updated at the current generation, individuals are selected at the same position at the current and previous generations. Here are three models for selecting individuals based on fixed methods: M-F1, M-F2, and M-F3. In M-R, individuals are selected in a random manner, which means that for individuals who need to be updated at the current generation, the location of individuals selected from the previous generation is random. Therefore, three models for selecting individuals based on random methods are M-R1, M-R2, and M-R3.

Suppose that the current generation is *t*, the position of the next generation individual is i , x_i^t is the i -th individual of the *t*-th generation, f_i^t is the fitness value corresponding to the *i*-th individual of the *t*-th generation, and *yi* is the individual generated by the basic algorithm. The corresponding fitness value is F^{t+1} . λ and μ_k are the weight vectors, satisfy, $\lambda > 0$, and $\mu_k > 0$, where the value of *k* is 1, 2, or 3, $\lambda + \Sigma \mu_k = 1$, $\lambda > 0$, and $\mu > 0$. *j_m* is a randomly selected individual position, *i* and *j* are not equal, and then, *m* is between 1 and the population size *N*. These six models can be defined as follows:

For the individual to be generated at the $(t + 1)$ -th generation, we select an individual from the random position *j* of the previous generation and combine with the individual at the current $(t + 1)$ -th generation to update individual at the next generation. The model is expressed as follows [\[28\]](#page-33-9):

$$
x_i^{t+1} = \lambda y_i^{t+1} + \mu x_j^t
$$

$$
\lambda = \frac{f_j^t}{F^{t+1} + f_j^t}
$$

$$
\mu = \frac{F^{t+1}}{F^{t+1} + f_j^t}
$$
(7)

when $i = j$, this model can be defined as M-F1.

• We randomly select an individual from the *t*-th generation and the (*t* - 1)-th generation. A total of two individuals are selected to update the next-generation individuals. The model is expressed as follows [\[44\]](#page-33-25):

$$
x_{i}^{t+1} = \lambda y_{i}^{t+1} + \mu_{1} x_{j_{1}}^{t} + \mu_{2} x_{j_{2}}^{t-1}
$$

\n
$$
\lambda = \frac{1}{2} \cdot \frac{f_{j_{2}}^{t-1} + f_{j_{1}}^{t}}{F^{t+1} + f_{j_{1}}^{t} + f_{j_{2}}^{t-1}}
$$

\n
$$
\mu_{1} = \frac{1}{2} \cdot \frac{F^{t+1} + f_{j_{2}}^{t-1}}{F^{t+1} + f_{j_{1}}^{t} + f_{j_{2}}^{t-1}}
$$

\n
$$
\mu_{2} = \frac{1}{2} \cdot \frac{F^{t+1} + f_{j_{1}}^{t}}{F^{t+1} + f_{j_{1}}^{t} + f_{j_{2}}^{t-1}}
$$
(8)

where $x_{j_1}^t$ and $x_{j_2}^{t-1}$ are individuals selected from random positions at the *t*-th and (*t* - 1)-th generations. The corresponding fitness function values are $f_{j_1}^t$ and f_{i2}^{t-1} . When $i = j_1 = j_2$, this model can be defined as M-F2.

• We randomly select three individuals from the *t*-th, (*t* - 1)-th, and (*t* - 2)-th generations to update individuals at the next generation. The model is expressed as follows [\[45\]](#page-33-26):

$$
x_{i}^{t+1} = \lambda y_{i}^{t+1} + \mu_{1} x_{j_{1}}^{t} + \mu_{2} x_{j_{2}}^{t-1} + \mu_{3} x_{j_{3}}^{t-2}
$$
\n
$$
\lambda = \frac{1}{3} \cdot \frac{f_{j_{3}}^{t-2} + f_{j_{2}}^{t-1} + f_{j_{1}}^{t}}{F^{t+1} + f_{j_{1}}^{t} + f_{j_{2}}^{t-1} + f_{j_{3}}^{t-2}}
$$
\n
$$
\mu_{1} = \frac{1}{3} \cdot \frac{F^{t+1} + f_{j_{3}}^{t-2} + f_{j_{2}}^{t-1}}{F^{t+1} + f_{j_{1}}^{t} + f_{j_{2}}^{t-1} + f_{j_{3}}^{t-2}}
$$
\n
$$
\mu_{2} = \frac{1}{3} \cdot \frac{F^{t+1} + f_{j_{3}}^{t-2} + f_{j_{1}}^{t}}{F^{t+1} + f_{j_{1}}^{t} + f_{j_{2}}^{t-1} + f_{j_{3}}^{t-2}}
$$
\n
$$
\mu_{3} = \frac{1}{3} \cdot \frac{f_{j_{2}}^{t-1} + f_{j_{1}}^{t} + F^{t+1}}{F^{t+1} + f_{j_{1}}^{t} + f_{j_{2}}^{t-1} + f_{j_{3}}^{t-2}}
$$
\n(9)

where $x_{j_1}^t$, $x_{j_2}^{t-1}$, and $x_{j_3}^{t-2}$ are individuals selected from random positions at the *t*-th, (*t* - 1)-th and (*t* - 2)-th generations. The corresponding fitness function values are $f_{j_1}^t$, $f_{j_2}^{t-1}$, and $f_{j_3}^{t-2}$. When $i = j_1 = j_2 = j_3$, this model can be defined as M-F3.

3.4 Motivation

As mentioned earlier, MOEAs have difficulty in dealing with MaOPs. The main reason is that a sharp increase in the number of non-dominated solutions in high-dimensional objective spaces can lead to greater convergence pressure; so, traditional MOEAs tend to stagnate in terms of convergence. Moreover, with the explosion of dimensions, the order of magnitude of MaOPs has become large. Although many researchers have tried to adopt MOEAs such as NSGA-II and MOEA/D to solve MaOPs, experiments show that owing to the curse of dimensions, MOEAs cannot balance the relationship between convergence and diversity. Recently, many studies have been conducted to modify the dominance relationship so as to enhance the ability to distinguish solutions [\[29](#page-33-10)[–31\]](#page-33-12). Convergence is an important property of solutions, and many researchers have used this property to make distinguishing solutions easier by improving a specific stage of the evolutionary process. In this improvement, the excellent solutions are selected one by one, thus increasing the selection pressure toward the Pareto front. The convergence indicator used in the 1by1EA is driven by this idea, which logically resembles the aggregate function in MOEA/D, but without a pre-defined weight vector. MOEA/D can optimize *N* standard quantum problems simultaneously, rather than by directly solving MOPs as a whole. The same idea applies to solving MaOPs, with 1by1EA being an example.

However, if only the effect of convergence is considered, the algorithm may be more likely to fall into local optimization. To solve this problem, the distribution of solutions in MaOPs also needs to be considered; generally, density estimation and niche techniques are more popular methods. The density estimation method estimates the neighborhood density value for each individual, selects individuals with smaller density value in the next evolution, and then deletes individuals with a larger density value. The characteristic of niche techniques is to form several stable sub-groups, namely niches, and then let individuals evolve in a specific environment. Jiang et al. [\[48\]](#page-33-29) improved SPEA by introducing an efficient reference direction-based density estimator, which can maintain the distribution of the population. The series of NSGA proposed by Deb et al. is a concrete embodiment of this technology, but the difficulty lies in determining the scope of the niche. Motivated by the above ideas, a distribution threshold is defined in the 1by1EA. Once a solution is selected by the convergence indicator, all solutions with distance to the selected one is less than the threshold are abandoned.

Based on the above description, 1by1EA can effectively balance the convergence and diversity of the solution, and it is theoretically feasible to select this algorithm for research in this paper. But this algorithm still has the following two limitations:

- 1by1EA has a certain absoluteness in the process of iteration, that is, to retain the best individuals at each iteration and discard all remaining individuals.
- 1by1EA does not perform as well when dealing with large-scale problems.

One of the main reasons why 1by1EA has difficulty in solving large-scale many-objective problems is that it only considers the relationship between different solutions. In the one-by-one selection strategy of the algorithm, once a solution is selected, all solutions with distribution distance from the selected solution less than the specified threshold are not valued. When the number of decision variables increases significantly, although this method can select the solutions one by one according to the convergence and distribution indicators, because the distribution of the solutions is very random, the effect of removing a lot of solutions within the threshold range of each selected solution may still be relatively large.

Therefore, to resolve the issues discussed above, this paper uses IFM to enhance 1by1EA and applies the enhanced algorithm to solve large-scale MaOPs. We preserve part of the historical information in the population in a fixed or random way, so that the transfer of information between different generations promotes the inheritance of population diversity by 1by1EA. While the individuals that are preserved are not necessarily the best, they can be used for the next renewal of the population, which allows for the continuous use of much useful information. Furthermore, this inheritance is of great significance for the solution set approaching the Pareto optimal front of large-scale MaOPs. After theoretical analysis, follow-up experiments also verified our idea. Thus, while ensuring the good coverage of Pareto optimal front, 1by1EA-IFM enhances the ability of the algorithm to solve large-scale problems. In addition, given the NSGA-III-IFM and MOEA/D-IFM proposed by Gu et al. [\[44\]](#page-33-25) and Zhang et al. [\[45\]](#page-33-26), the role of IFM is more effectively reflected in our research results than in previous research. The similarity between these several algorithms is that they all adopt the information retention mechanism of IFM, the difference is that our proposed 1by1EA-IFM is more convincing in the selection of candidate solutions. The details of 1by1EA-IFM are introduced in the next section.

4 1by1EA-IFM

In this section, in order to save the information in the historical iteration, we integrated IFM into 1by1EA. First, we introduce the overall framework of the proposed 1by1EA-IFM, and then explain the details. Finally, we choose one from the six algorithms as an example to describe the execution flow of the new algorithm in detail.

4.1 General framework of 1by1EA-IFM

To use the historical information in the previous iteration, we introduce the information feedback mechanism after producing new offspring individuals by the mutation operator before environmental selection and then updating these individuals. The approximate operation process of the algorithm based on 1by1EA is as follows. First, initialize the entire population and ideal points randomly, and then update the individuals in the population; this mainly includes three steps: parent mating selection to produce new offspring, operation operators for variation operations, and IFM to update individuals after variation. Finally, in the environment selection stage, the next generation parent population (*P*) is selected according to the one-by-one selection strategy. For more details about the algorithm, see Section [4.3.](#page-7-1)

4.2 Operators

A binary tournament selection strategy is proposed in mating selection. Density estimation based on distribution index and information ranking based on one-by-one selection are used as tournament strategies. Promising individuals can then be produced to form a mating pool. First, two individuals are randomly selected from the parent population. The lower ranked individual is selected from the two individuals, whose ranking value $rank(x)$ can be obtained from the one-by-one selection results of previous generations. If the two individuals are equally ranked, individuals with lower density estimates tend to be preferred, and density values can be calculated in conjunction with (5) and (6) according to the following formula [\[27\]](#page-33-8):

$$
d_k(x) = \frac{1}{\sum_{i=1}^k d_i^{min}(x) + 1}
$$
 (10)

where $d_i^{min}(x)$, $i = 1, ..., k$ is one of the *k* smallest values in $\{d_1(x), \ldots, d_{|k|}(x)\}$. Finally, if no amount of density estimation can distinguish between the two individuals, then a random selection is made between them.

The crossover operator used in this algorithm is a simulated binary crossover, which inherits useful genetic information from two or more offspring in the form of a single point crossover [\[49\]](#page-33-30). In this operator, p_1 and p_2 are two parents, c_1 and c_2 are two children, they satisfy $(p_1 + p_2)/2 = (c_1 + c_2)/2$. *β* is a spread factor, which is defined as $|(c_2 - c_1)/(p_2 - p_1)|$. If there is a random number ε in the range [0,1], then the range of value of ε will determine which of the following formulas will solve for *β*.

$$
c_1 = (p_1 + p_2) - 0.5\beta(p_2 - p_1)
$$

\n
$$
c_2 = (p_1 + p_2) + 0.5\beta(p_2 - p_1)
$$

\n
$$
c(\beta) = \begin{cases} 0.5(\varepsilon + 1)\beta^{\varepsilon}, & \varepsilon \le 0.5\\ 0.5(\varepsilon + 1)\frac{1}{\beta^{\varepsilon+2}}, & \varepsilon > 0.5 \end{cases}
$$
\n(11)

After the crossover operator is executed, the individual is updated by the polynomial mutation [\[50\]](#page-33-31). In this operator, p_m is defined as a parent, and q_m is the offspring produced by mutation operator. σ is a random number in the range [0,1]. *η* is a distribution index. u_m and l_m are the upper and lower bounds of *m*-dimensional decision variables. Thus,

the form of this mutation operator is $q_m = p_m + \delta(u_m - l_m)$. The following formula can be referred to calculate *δ*:

$$
\delta = \begin{cases}\n[2\sigma + (1 - 2\sigma)(1 - \delta_1)^{\frac{1}{\eta + 1}}] - 1, & \sigma \le 0.5 \\
1 - [2(1 - \sigma) + 2(\sigma - 0.5)(1 - \delta_2)^{\eta + 1}]^{\frac{1}{\eta + 1}}, & \sigma > 0.5\n\end{cases}
$$
\n
$$
\delta_1 = (p_m - l_m)/(u_m - l_m)
$$
\n
$$
\delta_2 = (u_m - p_m)/(u_m - l_m)
$$
\n(12)

4.3 1by1EA-F1

Here, we describe in detail how 1by1EA is combined with IFM and give detailed workflow. We mainly chose the model of selecting an individual in a fixed way from the previous generation as an example. The new algorithm after their combination was named 1by1EA-F1. The remaining five algorithms (1by1EA-F2, 1by1EA-F3, 1by1EA-R1, 1by1EA-R2, and 1by1EA-R3) all performed the same as this algorithm. The main execution of 1by1EA-F1 can be interpreted as the following process:

- Step 1 Initialization. Generate random population *P* and ideal point *Zmin*. Rank of each solution in one-by-one selection and initialize distribute threshold.
- Step 2 Update. Supposing the current generation is *t* and $i = 0$ (*i* is the subscript of the individual).
- Step 2.1 Generate offspring Y^{t+1} by using the crossover and mutation operator of the original 1by1EA. y_i^{t+1} produced by the above process and constitute Y^{t+1} , and then, their fitness value will be calculated by functions.
- Step 2.2 Calculate x_i^{t+1} by combining y^{t+1} with x_i^t according to [\(7\)](#page-4-2). K^t is the set of x_i^{t+1} .
- Step 2.3 Determine whether the calculation has been completed for each individual. If it is satisfied, perform Step 2.4. If not, $i = i + 1$ and repeat Step 2.1 and Step 2.2. Step 2.4 Environmental selection.
- Step 2.4.1 Normalize each solution in K^t .
- Step 2.4.2 Calculate the convergence indicator $c(x)$ and distribution indicator $d(x)$ of each solution $x \in K^t$, select some excellent one by one and put them into K_s .
- Step 2.4.3 The corner solution is chosen to form set K_s .
- Step 2.4.4 Guarantee the number of individuals in K_s to be *N*, which is the population size.
- Step 2.4.5 Update the distribution threshold.
- Step 2.4.6 Verify whether termination condition is met. The termination condition is the maximum number of iterations defined in advance. If it is satisfied, go to Step 3. If not, $t = t + 1$ and repeat Step 2.

Step 3 Output. Output P^{t+1} to be equal to K_s .

The flow chart of 1by1EA-F1 can be shown in Fig. [1.](#page-8-0)

4.4 Complexity

In this subsection, the complexity of the algorithm is analyzed. For convenience, assuming that the problem to be optimized has *m* objectives and *n* decision variables, the decision-maker requires *n* solutions. Its complexity is analyzed according to Algorithm 1 and the above steps. In addition, it is assumed that the size of the population is set to *N*, and the generation is set to be *t*. Therefore, in the initialization process (Step 1), the computational complexity of generating the ideal point is O(*mN*). This is because in the process of generating the ideal point, an aggregation function is used to aggregate all the objectives, and each component in the ideal point corresponds to the minimum value of each sub-objective. Furthermore, generating a random population and ranking of each solution require O(*nN*) and $O(mN)$ computations, respectively. As the convergence indicator and the distribution indicator need to be calculated according to the ideal point and the solution needed by the decision maker, the computational complexity in Step 2.4.2 is $O(mN^2)$ and $O(nN^2)$, respectively. Then, generation of offspring (Step 2.1) requires $O(\frac{N}{2}(n + n))$, because the size of the mating pool is set to *N*. In Step 2.2 and Step 2.3, IFM begins to work, because each generation in the iterative process needs information from the historical generation, the computational complexity is about $O(\frac{N}{2}nmt)$. During environmental selection, owing to the introduction of oneby-one selection strategy, the best case of computational complexity is $O(N^2)$, while the worst case is $O(N^3)$, and *N* individuals are linearly assignment to the ideal points. Furthermore, it is fact that generally $N \gg m$ and $N \gg n$ in MaOPs. Therefore, when the iteration termination condition is reached, the complexity of the proposed algorithm is $O(tN^3)$.

5 Experiments

In this section, we describe the verification process of the proposed 1by1EA-IFM in detail. The experiment was divided into three stages. First, we compared the performance of the proposed 1by1EA-IFM with 1by1EA in dealing with large-scale multi-objective test problems (LSMOPs) with different decision variables. Then we focused on the work of previous researchers based on IFM and compared the outstanding algorithms in 1by1EA-IFM (selected in the first stage) with the excellent algorithms in NSGA-III-IFM and MOEA/D-IFM on the same indicators and test issues. In addition, we focused on MOKP and selected two other comparison algorithms to obtain the verification results. In the last stage, we chose an algorithm specifically designed to solve the large-scale problem and

the other five state-art-of MaOEAs to measure our work from a vertical perspective. In order to avoid contingency, all algorithms had been verified to deal with LSMOPs [\[51\]](#page-33-32) with different numbers of decision variables. In these problems, the parameters that had to be set included the number of objectives (*m*), the dimension of the decision variable (*n*), and the number of subcomponents in each variable group (n_k) . Here, we took $n = m * 100$, $n_k = 5$. These settings proved valid when LSMOPs were proposed as test functions [\[51\]](#page-33-32).

As a test suite, LSMOPs were composed of optimization problems of linear problems, convex problems, multimodel problems, and the disconnected Pareto optimal front, which were defined by specified functions. The different characteristics of these fronts are shown in Table [1.](#page-9-0) In this paper, the objective number *m* did not exceed 15, and most experiments used *m* values of 3, 5, 8, 10, and 15. The inverted generational distance (IGD) [\[52\]](#page-33-33) and the generational distance (GD) [\[53\]](#page-33-34) were adopted as performance indicators in most experiments. GD was used to evaluate the convergence of the algorithm. It was mainly used to express the separation distance between the approximate PF and the real PF. Its calculation formula is as follows $[54]$:

$$
GD = \frac{\left(\sum_{i=1}^{n} d_i^m\right)^{\frac{1}{m}}}{n} \tag{13}
$$

where *n* is the number of points on the approximate PF, *m* is the number of dimensions of the objective, and *di* represents the average value of the nearest Euclidean distance between each point on the approximate PF and the reference point on the true PF. The smaller the value, the better the convergence of the algorithm.

However, as GD can only evaluate the convergence of the algorithm, in order to simultaneously evaluate the convergence and diversity of the algorithm, IGD was used, it is the solution obtained by mapping the reference point on the real PF to the algorithm. The main idea was to take points uniformly from the real PF, find the closest point on the known PF to the point on the real PF for each point on the real PF, and then add the distances between these points

Table 1 The characteristics of LSMOPs

Problem	Characteristics	Modality
LSMOP1	Linear	Unimodal
LSMOP2	Linear	Mixed
LSMOP3	Linear	Multi-modal
LSMOP4	Linear	Mixed
LSMOP5	Convex	Unimodal
LSMOP6	Convex	Mixed
LSMOP7	Convex	Multi-modal
LSMOP8	Convex	Mixed
LSMOP9	Disconnected	Mixed

and average them. The calculation formula is as follows [\[55\]](#page-33-36):

$$
IGD = \frac{\sum_{i=1}^{n} |d_i|}{n}
$$
\n(14)

where n is the number of points on the true PF. The smaller the value is, the better the overall performance of the algorithm is.

All experiments in this paper were implemented in Matlab on PlatEMO with two Intel G3204 CPUs. Considering the characteristics of evolutionary algorithms, the IGD and GD values shown in all the following tables were the results of indicators after 20 independent iterations on specific problems, taking the average and standard deviation. More importantly, in terms of statistical method, the Mann-Whitney-Wilcoxon rank-sum test [\[56\]](#page-33-37) was used to determine whether there is a statistically significant difference in indicators between one algorithm and another algorithm, and the null hypothesis was rejected at a significance level of 5%. Note that the datasets on which statistical tests relied contain a set of values that the algorithm calculates indicators for on the LSMOPs' test suite with different objectives. All the above settings were used in all subsequent experiments.

5.1 Comparison of 1by1EA with 1by1EA-IFM

5.1.1 Parameter settings

For the fairness of the results, in the simulated binary crossover and polynomial mutation, the two distribution index values were set to 20. The probability of crossover and mutation were 1.0 and 1/*n*, respectively, where *n* was the number of decision variables. The parameter settings of these evolutionary operator were the same as those in recent studies, and these settings proved to be able to generate good offspring in [\[34,](#page-33-15) [42,](#page-33-23) [45,](#page-33-26) [57,](#page-34-0) [58\]](#page-34-1). From the historical research on evolutionary algorithms, we know that when the above parameters are set to other values, the results are more insensitive to the parameters; so, the values of these parameters were selected in this work to mitigate the impact of parameter settings on the results. The maximum number of iterations of the algorithm was 10000, and their population sizes were the same. In all algorithms, the value of *k* in [\(10\)](#page-6-1) was set to 1/*N* to balance the computational cost and accuracy in density estimation, and the threshold of the number of individuals (*R*) selected in advance was set to 1 to enable the population to balance the distribution and convergence of the population. The setting of these two experimental parameters in the original algorithm was proved in their own experiments. When the threshold was 1, the convergence and diversity of the solution could be properly balanced. The population size was set from 100 to 200 according to the different dimensions of the objective space.

5.1.2 Experimental results and analysis

Table [2](#page-11-0) lists the average and standard deviation of the IGD values for each algorithm running on LSMOP1-9 with 3, 5, 8, 10, and 15 objectives. The best result in each row of this table is bolded, which applies in all the similar rest of tables below. The symbols "+" and "-" indicate whether the null hypothesis was accepted by the Mann-Whitney-Wilcoxon rank-sum test at the significance of 5% level, that is, whether the IGD and GD results of 1by1EA-IFM were better or worse than those of 1by1EA. Further, the symbol "=" indicates that the corresponding null hypothesis was rejected. Note that the number of "+", "-", and "=" in the last row respectively indicates how many times 1by1EA-IFM was obviously better, obviously worse, and similar to 1by1EA over each test problem on different objectives. The best results in each row of the table are emphasized.

In all 45 test instances, the sum of the best results achieved by the six algorithms in 1by1EA-IFM was 32, while the best results of 1by1EA could only reach 13. In the test example with a lower objective of less than 10 targets, 1by1EA-F1 and 1by1EA-F2 performed extremely well. The sum of the best results obtained by the two was 15, which accounts for almost 50% of the number of best results. Among the three-objective test problems, 1by1EA-F1 could obtain the best results of the average IGD value on LSMOP1/3/5/6/8/9, and the other three best results were respectively obtained by 1by1EA-R1 and 1by1EA-F3. Among the five-objective test functions, 1by1EA-F1 and 1by1EA-R1 could obtain the best results of six out of LSMOP1-9. Among the eight-objective test functions, 1by1EA-F1 and 1by1EA-F2 had a similar performance. In the 10- and 15-objective test functions, the best results were obtained more by 1by1EA-F2 and later algorithms. The last row of the table shows that the values that were superior to 1by1EA were 29, 26, 25, 22, 24, and 23, respectively. This shows that our proposed algorithm could beyoud 1by1EA on most test functions from the perspective of the number of objectives. From the level of the test suite, motivated by the IGD results, it can be seen from the last column in the table that the best results obtained by 1by1EA algorithm were concentrated on LSMOP2-4, whose PF were linear. Therefore, to some extent, it can be considered that 1by1EA is more suitable for the linear PF, and our proposed algorithm has wider application scope, which can be interpreted that the sampled points from the Utopian PF for the proposed algorithm are the Pareto-optimal solutions owing to the mixed feature of the PF, and the information retaining in IFM has good effects.

The findings obtained above are more obvious in the results of the GD value. As shown in Table [3,](#page-13-0) in all 45 test examples, 43 best results were achieved, and 1by1EA could only achieve two. Unlike the result obtained by the IGD value, the best result of the GD value was evenly distributed among the six algorithms proposed. As can be seen from the last row in the table, the sum of the numbers of different objective problems obtained by the six algorithms with significantly superior performance was 36, 38, 37, 31, 40, and 38 respectively. Compared with the IGD, GD could be further improved. As an indicator to consider the convergence, GD could not measure the quality of the solution set more comprehensively than IGD, but it could also reflect the performance of the algorithm. We previously mentioned that IFM can increase the diversity of solutions in the population at the next iteration through information feedback mechanism. The experimental results of the GD value confirm the conclusion that some information of the solution is retained the next time through the information reuse strategy, which can improve the convergence of the algorithm by increasing the diversity of the solution set. There is no relevant strategy in 1by1EA; so, 1by1EA-IFM had better convergence. Moreover, when the dimensionality of the problem was enhanced, when the decision variable increased to 1500, the performance of our proposed algorithm was not affected. Moreover, excellent results could be more evenly distributed in various LSMOPs with different characteristics, that is, the algorithm could effectively face complex PF, and 1by1EA can hardly solve diverse problems. This finding still relies on the role of information retention mechanisms; in complex environmental situations that are simulated, the continuous reuse of historical information can make the evolution of the population more suitable for environmental changes.

In order to show the results more clearly, here, we used 10-objective problem as an example to illustrate the performance of the six algorithms and the 1by1EA. Figure [2](#page-15-0) shows the performance of the seven algorithms on LSMOP1-9 when $m = 10$ along with the iteration. It can be clearly found that on LSMOP1/5/7/8/9, the index values of 1by1EA were all significantly higher than those of the proposed algorithm, which shows that in most LSMOPs, our algorithm was capable of solving these problems.

5.1.3 Statistical analysis of the results

We performed two statistical tests with the above obtained results. It should be emphasized that these two tests can further illustrate the superiority of our algorithms from a statistical point of view. As we knew, the results obtained

Fig. 2 The mean of IGD and GD values on LSMOP1-9 with 10 objectives

from our experiments were the mean and standard deviation of IGD and GD. We chose the standard deviation of metrics as the statistic to be measured because the magnitude of the standard deviation directly measures the degree of dispersion of the data, so that all data characteristics are adequately reflected rather than making the positive and negative extremes cancel each other out as in the case of the mean.

First, after integrating the data in tables before, Friedman's non-parametric test was applied to check whether there were difference among all algorithms. The results in Tables [2](#page-11-0) and [3](#page-13-0) are further analyzed using Friedman test, as shown in Table [4.](#page-16-0) Since the confidence interval is 95%, in a χ^2 distribution with 6 degrees of freedom, we gained a *p*-value of 0.000006, which was significantly below the significance level (α =0.05). The results of this statistical test showed that our proposed six algorithms were significantly different from 1by1EA. The best rank was obtained by 1by1EA-F1, however, 1by1EA reached the worst rank.

After obtaining the above results, Holm post-hoc test was performed to verify the statistical significance of the six new algorithms. Since the proposed 1byEA-F1 obtained the best ranking, it was considered as control algorithm, and by pairwise and multiple comparisons, we obtained the adjusted *p*-values after using Holm post-hoc procedure, as shown in Table [5.](#page-16-1) In this table, the algorithms were assigned an ordinal number *i* that indicated the degree of excellence of the algorithm, with the first column of the table representing from the best to the worst from top to bottom. The null hypothesis was rejected on the premise that the *p*-value obtained by Holm post-hoc procedure need to be less than the adjusted $\alpha(\alpha = \alpha/i)$. Therefore, we can conclude with a high degree of certainty that there is not significant difference among our proposed six algorithms, and they are significantly better than 1by1EA.

5.2 Comparison with NSGA-III-IFM and MOEA/D-IFM

In the first phase of the experiment, we verified the performance of 1by1EA-IFM. As there are many models,

Table 4 Friedman test ranking for 1by1EA-F1, 1by1EA-F2, 1by1EA-F3, 1by1EA-R1, 1by1EA-R2, 1by1EA-R3, and 1by1EA

Algorithm	Average ranking
$1by1EA-F1$	2.68
$1by1EA-F2$	2.97
$1by1EA-F3$	3.90
1by1EA-R1	4.25
1by1EA-R2	3.12
$1by1EA-R3$	3.75
1by1EA	5.10

Table 5 Results obtained using Holm's post-hoc test

i	Algorithm α/i		p -value of Holms Null hypothesis	
		1 lby1EA-F2 $0.05/1 = 0.0500$ 1.000		Not reject
		2 1by1EA-R2 $0.05/2 = 0.0250$ 1.000		Not reject
		3 1by1EA-R1 $0.05/3 = 0.0166$ 0.022		Not reject
		4 1by1EA-R3 $0.05/4 = 0.0125$ 0.013		Not reject
		5 1by 1EA-F3 $0.05/5 = 0.0100$ 0.012		Not reject
		$1 \text{by} 1 \text{EA}$ $0.05/6 = 0.0083$ 0.003		Reject

1by1EA-F1 used as control algorithm

here, we counted the sum of the number of GD and IGD indicators that were significantly better than the original algorithm in Experiment 1. As shown in Table [6,](#page-17-0) the best algorithm was 1by1EA-F1, because the final results of 1by1EA-F2 and 1by1EA-R2 were both 64. Thus, we chose these three algorithms for subsequent experiments.

5.2.1 Comparison of 1by1EA-IFM with NSGA-III-IFM

In order to avoid the contingency of the results, we also used the same combination as in the original paper. The rest of the parameter settings, the number of iterations, and indicator selection were also the same as those in Experiment 1. We chose the best two algorithms, NSGA-III-F1 and NSGA-III-R1, for comparison to reflect the accuracy of our results.

Table [7](#page-17-1) shows the statistical results from the Mann-Whitney-Wilcoxon rank test. Among all the best results highlighted, 1by1EA-IFM accounted for 39, and only six of them appeared in NSGA-III-IFM. It can be seen from Table [7](#page-17-1) that the number with better results by 1by1EA-F1, 1by1EA-F2, and 1by1EA-R2 on LSMOPs was 35, 25, and 23, respectively. 1by1EA-F1 still had the best performance among these five algorithms. 1by1EA-F2 was second, while 1by1EA-R2 did not seem to have much of an effect. Furthermore, as listed in Table [8,](#page-18-0) all the best results were obtained by our proposed algorithm. Unexpectedly, 1by1EA-R2 performed the best on the GD indicator, and the results with superior performance could cover all test example. The performance of the remaining two algorithms on this indicator was the same. Therefore, in general, the performance of 1by1EA-IFM was better. The introduction of the one-by-one selection strategy may have made 1by1EA-IFM superior to NSGA-III-IFM based on reference point selection in terms of convergence. In addition, when we examined the role of the same model, the results of this horizontal comparison were more likely to show the scalability of IFM. Although 1by1EA makes a significant effort to balance convergence and distribution, the retention of information can further improve the utilization of

Table 6 The GD and IGD values

Metric	M	\boldsymbol{N}	1 by $1EA-F1$	1 by $1EA-F2$	1by1EA-F3	1 by $1EA-R1$	1by1EA-R2	1by1EA-R3
GD	3	300			6	Ω	Q	
		500	6	n				
	8	800	6					
	10	1000	8					
	15	1500	9					
IGD	3	300			6	0		
		500						
	8	800	6					
	10	1000	6					
	15	1500	5					
Total			65	64	62	53	64	61

Table 7 IGD values of 1by1EA-F1, 1by1EA-F2, 1by1EA-R2, NSGA-III-F1, and NSGA-III-R1 on LSMOP1-9 with 3, 5, 8, 10, and 15 objectives

Problem	M N		$1by1EA-F1$	$1by1EA-F2$	$1by1E A-R2$	NSGA-III-F1	NSGA-III-R1
	5	500	$2.5773e+0$ (8.27e-2) +	$4.1687e+1(1.02e+2) =$	$3.0486e+1(6.40e+1)$ -	$2.9911e+0(1.01e-1) =$	$2.9614e+0(7.70e-2)$
	8	800	$9.2911e+1(2.47e+1) +$	$2.2766e+2(3.78e+1) =$	$2.7649e+2(4.60e+1)$ -	$2.1632e+2(5.98e+1) =$	$2.1857e+2(7.20e+1)$
	10	1000	$1.4373e+2(2.62e+1) =$	$2.6070e+2(3.51e+1)$ -	$2.7778e+2(3.84e+1)$ -	$1.8380e+2(5.80e+1) =$	$1.7117e+2(5.13e+1)$
	15	1500	$1.7255e+0$ (1.26e-2) +	$7.6932e+1(8.02e+1)$ -	$1.1294e+2(8.30e+1)$ -	$1.7629e+0$ (1.19e-2) =	$1.4130e+1(5.53e+1)$
LSMOP8 3		300	$5.5532e-1$ $(4.53e-3) +$	$5.5460e-1(2.76e-3) +$	$5.5830e-1$ $(3.88e-3) +$	$9.7049e-1(1.63e-2) +$	$9.7692e-1(9.28e-3)$
	5	500	$9.3379e-1(2.97e-2) +$	$9.2392e-1(1.66e-2) +$	$9.3039e-1(1.57e-2) +$	$1.1704e+0$ (5.29e-2) =	$1.1809e+0(3.82e-2)$
	8	800	$1.1235e+0$ (2.43e-2) +	$1.2413e+0$ (4.84e-2) +	$1.3282e+0$ (4.81e-2) +	$1.8535e+0$ (4.39e-1) =	$1.9196e+0(3.58e-1)$
	10	1000	$1.2435e+0$ (4.00e-2) +	$1.2981e+0(4.41e-2) +$	$1.3400e+0$ (5.73e-2) +	$1.8709e+0(2.63e-1) =$	$1.9739e+0(2.14e-1)$
	15	1500	$1.2713e+0$ (2.41e-2) +	$1.2770e+0$ (1.80e-2) +	$1.2824e+0(3.73e-2) +$	$1.3047e+0$ (1.38e-2) =	$1.3023e+0(1.75e-2)$
LSMOP9 3		300	$1.3354e+1(8.07e-1) +$	$1.4615e+1$ (9.20e-1) +	$1.5528e+1(1.14e+0) +$	$3.0398e+1(3.97e+0) =$	$2.9282e+1(5.20e+0)$
	5	500	$3.8021e+1(1.96e+0) +$	$4.0998e+1(1.68e+0)+$	$4.3014e+1(1.48e+0) +$	$1.0937e+2(1.42e+1) =$	$1.1027e+2(1.41e+1)$
	8	800	$7.1140e+1(2.18e+0) +$	$7.2537e+1(2.38e+0)+$		7.4044e+1 $(2.51e+0)$ + 2.5767e+2 $(4.68e+1)$ = 2.6664e+2 $(4.12e+1)$	
	10	1000	$1.0992e+2(5.01e+0)+$	$1.0806e+2(2.94e+0)+$	$1.1229e+2(4.63e+0)+$	$6.0376e+2(9.30e+1) = 5.9636e+2(1.89e+2)$	
	15	1500			$4.1853e+2$ (2.22e+1) + $3.9874e+2$ (1.20e+1) + $4.0824e+2$ (1.44e+1) + 2.2210e+3 (7.61e+2) = 2.1716e+3 (8.94e+2)		
$+/-/=$			35/8/2	25/13/7	23/17/5	2/1/42	

Table 8 GD values of 1by1EA-F1, 1by1EA-F2, 1by1EA-R2, NSGA-III-F1, and NSGA-III-R1 on LSMOP1-9 with 3, 5, 8, 10, and 15 objectives

Table 8 (continued)

solutions in populations. As a general information feedback framework, historical research can guide us to further optimize the evolutionary direction of algorithms.

Similarly, in order to more clearly illustrate the results, we took LSMOP5 as an example to show the changes in the IGD and GD values on different objectives with iterations. Figure [3](#page-20-0) shows the performance of the five algorithms related to this section. Our algorithm had the best overall performance for all high-dimensional problems. The average value of LSMOP5 in 1by1EA-IFM was lower than that of NSGA-III-IFM. One research focuse of this work was to test whether the algorithm has the ability to solve large-scale MaOPs, and the high-dimensional results were very promising. The results of the comparison of our proposed algorithm with NSGA-III-IFM verify that our algorithm has strong competitiveness in solving the same large-scale MaOPs.

5.2.2 Comparison of 1by1EA-IFM with MOEA/D-IFM

In the experiment at this stage, the value of the objective number *m* was 5, 10, and 15. The test problem was still LSMOPs. The other parameters in our algorithm were the same as those used in previous experiments. We also used a HyperVolume (HV) [\[59\]](#page-34-2) indicator to evaluate the results. It measures the size of the area volume value in the target space enclosed by the non-dominated solution set and the reference point obtained by the algorithm. The larger the value, the better the overall performance of the algorithm that was considered. As the best one, MOEA/D-R1 was selected for comparison.

Table [9](#page-19-0) shows the best results in each row with emphasis. From all the statistics obtained from the Mann-Whitney-Wilcoxon rank test with the best IGD indicator values, our proposed algorithm occupied 21 of the 27 test examples, while MOEA/D-R1 only occupied six. In particular, 14 were obtained by 1by1EA-F1, which is relatively high, accounting for almost 50% of all test examples. Combined with the HV results shown in Table [10,](#page-22-0) we could still make this observation. Here, we only provided statistics on the number of HV results that were better than the MOEA/D-R1. From the table, in terms of all objective test problems, 1by1EA-F1 had the best performance, followed by 1by1EA-F2. For the nine test questions with three, five, and 10 objectives, the number of excellent results that the 1by1EA-IFM algorithm could achieve was eight, six, and nine, respectively. Similarly, the results of 1by1EA-F1 in these dimensions were seven, five, and seven, respectively. Therefore, the comprehensive performance of our proposed 1by1EA-IFM was better than that of MOEA/D-IFM. MOEA/D-IFM uses a predefined objective search to solve the problem of a slow search process in a high-dimensional space. Individual selection and retention in IFM improve the sensitivity of predefined targets, especially in a largescale high-dimensional space. For LSMOP with different

Fig. 3 The mean of IGD and GD values on LSMOP5 with 3, 5, 8, 10, and 15 objectives

fronts, 1by1EA-IFM could solve linear problems more effectively than MOEA/D-IFM, which may have been caused by target decomposition in MOEA/D. However, for the high-dimensional problems of 15 objectives, MOEA/D-IFM could also play a better role; so, the enhancement of the connection between solutions in IFM effectively improved the performance of the evolutionary algorithm.

In order to demonstrate the ability of our proposed algorithm to solve higher-dimensional problems, here, we took LSMOP1-9 with 15 objectives as an example to show the changes in the IGD values obtained by the

Table 9 IGD values of 1by1EA-F1, 1by1EA-F2, 1by1EA-R2, and MOEA/D-R1 on LSMOP1-9 with 5, 10, and 15 objectives m I SMOP1-9 with 5 10 and MOFA/D-R1 1h v $1 \text{F} \Delta \text{F}$ 1 by $1FA$ _{-F}? of 1 by $1F \Delta - F1$

\boldsymbol{M}		1 by $1EA$ -F 1	1 by $1E$ A-F2	1 by $1EA-R2$	MOEA/D-R1
5	500				
10	1000	$\overline{}$			
15	1500				

Table 10 The optimal HV result of 1by1EA-F1, 1by1EA-F2, 1by1EA-R2, and MOEA/D-R1 on LSMOP1-9 with 5, 10, and 15 objectives

above algorithm as it iterated. As shown in Fig. [4,](#page-22-1) on LSMOP2/3/4/7/8, the performance of 1by1EA-IFM was significantly better than that of MOEA/D-R1, that is, in the nine test functions, for more than half of the problems, the algorithms we proposed all outperformed those in previous research. However, it is undeniable that, MOEA/D-IFM was more suitable for solving problems characterized by disconnected of PF, as the results of LSMOP9 demonstrate.

5.2.3 Statistical analysis of the results

Based on the IGD and GD values obtained by all above comparison, we performed the same statistical analysis in two steps. As shown in Tables [11](#page-23-0) and [12,](#page-23-1) the computed ranking by using Friedman's test for the above two comparison were displayed respectively. In two *χ*² distributions with 4 and 3 degrees of freedom respectively,

Fig. 4 The mean of IGD values on LSMOP1-9 with 15 objectives

Table 11 Friedman test ranking for 1by1EA-F1, 1by1EA-F2, 1by1EA-R2, NSGA-III-F1, and NSGA-III-R1

Algorithm	Average ranking
$1by1EA-F1$	2.45
$1by1EA-F2$	2.59
$1by1EA-R2$	2.70
NSGA-III-F1	3.69
NSGA-III-R1	3.56

these average ranking demonstrated that 1by1EA-IFM was different from NSGA-III and MOEA/D-IFM with a *p*-value $= 0.000000008$ under the significance level $\alpha = 0.05$. The best ranking was obtained by 1by1EA-IFM, and the performance of our proposed algorithms were close to and generally better than NSGA-III-IFM and MOEA/D-IFM. These ranking were helpful for us to determine the order number in Holm's test.

When we achieved Friedman's test, Holm's method was used to compare the best ranking algorithm with the remaing algorithms. Since the proposed 1byEA-F1 obtained the best ranking, it was considered as control algorithm. Tables [13](#page-23-2) and [14](#page-23-3) respectively displayed the adjusted *p*values by Holm post-hoc procedure based on above two experiments. The null hypothesis was rejected when the achieved *p*-value lower than the adjusted α ($\alpha = \alpha/i$). After analyzing, all the *p*-values about NSGA-III-IFM and MOEA/D-IFM obtained by Holm post-hoc were lower than $α$. Thus, 1by1EA-F1 is statistically better than those algorithms in NSGA-III-IFM and MOEA/D-IFM.

5.2.4 Multi-objective knapsack problem (MOKP)

Here, we adopted MOKP and selected two recently proposed algorithms, AGE-MOEA and PREA, to compare with the results of 1by1EA-IFM and MOEA/D-IFM on this problem.

AGE-MOEA [\[60\]](#page-34-3) comprehensively considers the diversity of individuals in the population and the proximity to the ideal point. It does not have a pre-supposed PF, but uses a quick geometry program to estimate PF. The increase in computational complexity is more obvious.

Table 12 Friedman test ranking for 1by1EA-F1, 1by1EA-F2, 1by1EA-R2, and MOEA/D-R1

Algorithm	Average ranking
1 by $1EA-F1$	1.78
1 by $1EA-F2$	2.41
1 by $1EA-R2$	1.92
MOEA/D-R1	3.89

Table 13 Results obtained using Holm's post-hoc test

<i>i</i> Algorithm	α/i	p -value of Holms Null hypothesis	
	1 1by1EA-F2 $0.05/1 = 0.0500$ 1.000		Not reject
	2 1by1EA-R2 $0.05/2 = 0.0250$ 1.000		Not reject
	3 NSGA-III-R1 $0.05/3 = 0.0166$ 0.009		Reject
	4 NSGA-III-F1 $0.05/1 = 0.0100 0.004$		Reject

1by1EA-F1 used as control algorithm

PREA [\[61\]](#page-34-4) is a kind of MaOEA based on probability index. PREA introduces a ratio index with infinite norm so that the algorithm can select a promising area in the target space. Then, a population selection strategy based on parallel distance selection is used to offspring populations in this area.

In order to adapt to the problem studied in this paper, we set *M* to three, five, and eight, and *N* was set to 300, 500, and 700, respectively. The number of populations and the maximum number of iterations remained unchanged. The weight was a random number between 10 and 100. The experimental results are shown in Table [15.](#page-24-0) It can be seen that the average results of IGD values obtained by 1by1EA-IFM were significantly better than those of other MaOPs. The best part compared with other comparison algorithms is highlighted in the table. 1by1EA-F2 could obtain the optimal IGD value in the problem of 3 backpacks with 300 items and three backpacks with 500 items. 1by1EA-R2 could reach the smallest IGD value in the remaining problems. This result not only proves the advantages of 1by1EA-IFM in solving MOKP, but also shows that its performance was superior to MOEA/D-IFM.

5.3 Comparision with other many-objective algorithms

5.3.1 Compared algorithms

In order to evaluate the performance of the proposed new algorithm, we selected the best performing algorithm among the six proposed new algorithms and compared it with the following six advanced evolutionary algorithms proposed

Table 14 Results obtained using Holm's post-hoc test

i Algorithm	α/i	p -value of Holms Null hypothesis	
	1 1by1EA-R2 $0.05/1 = 0.0500$ 1.000		Not reject
	2 1by 1EA-F2 $0.05/2 = 0.0250$ 0.439		Not reject
	3 MOEA/D-R1 $0.05/3 = 0.0166$ 0.002		Reject

1by1EA-F1 used as control algorithm

able 15 IGD values of 1by1EA-F1, 1by1EA-F2, 1by1EA-R2, 1by1EA, AGEMOEA, PREA and MOEA/D-R1 on MOKP **Table 15** IGD values of 1by1EA-F1, 1by1EA-F2, 1by1EA-R2, 1by1EA, AGEMOEA, PREA and MOEA/D-R1 on MOKP

in recent years on LSMOP1-9: MaOEA-IGD, NSGA-II-SDR, DGEA, DEA-GNG, LMOCSO, and MSEA. Based on Section [5.1,](#page-9-1) we know that the best algorithm was 1by1EA-F1, and this result could be verified in almost all subsequent experiments. The selected comparison algorithms almost cover the main categories of MaOEAs in Section [1,](#page-0-0) and there are algorithms specifically for solving large-scale MaOPs.

- MaOEA-IGD [\[34\]](#page-33-15) is a many-objective evolutionary algorithm based on the IGD indicator. Specifically, it uses the IGD index at each generation and adopts a computationally efficient advantage comparison method through a linear allocation mechanism from a global perspective to choose a solution with better convergence and diversity. As the algorithm was evaluated based on the IGD index in the whole process and IGD was included in all the experimental indexes in this paper, this algorithm was selected as the comparison algorithm.
- A new algorithm based on Pareto dominance relation NSGA-II-SDR [\[62\]](#page-34-5) is proposed. It is an improvement of NSGA-II, using an adaptive niche technology based on the angle between candidate solutions to make only the convergent optimal candidate solution in the niche be the non-dominated solution. This algorithm has shown effectiveness in solving MaOPs.
- DGEA [\[42\]](#page-33-23) is an algorithm specifically proposed to solve LSMOPs. It uses an adaptive offspring generation method to generate promising candidate solutions in high-dimensional spaces. However, the algorithm has not been verified in the dimension of many-objective problems. Based on this, we performed a comparison in high dimensions.
- DEA-GNG [\[63\]](#page-34-6) is a decomposition-based multiobjective evolutionary algorithm, which is guided by the growing neural gas network (GNG), and uses the nodes in the GNG as reference vectors according to an adaptive strategy. The method proposed by this algorithm is relatively novel; so we chose this algorithm as a comparison algorithm.
- LMOCSO [\[64\]](#page-34-7) combines a competitive swarm optimizer (CSO)-based with LSMOPs, adopting a new particle updating strategy that suggests a two-stage strategy to update position, so that search efficiency is improved. Similar to DGEA, the method is also used to solve LSMOPs, but when the number of objectives is increased, the performance of LMOCSO has not been proven, which is why we chose it as a comparison method.
- As a newly proposed algorithm, MSEA [\[65\]](#page-34-8) also allows populations to evolve toward PF through diversity preservation. This method divides the optimization

process into multiple stages, and updates the population with different selection strategies at different stages. Since MSEA is also based on diversity improvement algorithm, we selected it for comparison to reflect the performance of IFM for population diversity.

5.3.2 Parameter settings

All comparison algorithms used the following parameter settings. In simulated binary crossover and polynomial mutation, both distribution indexes were set to 20. The probability of crossover and mutation was 1.0 and 1/*n*, respectively, where *n* is the number of decision variables. The reason for these settings is discussed in Section [5.1,](#page-9-1) and the fixed setting could highlight the robustness and generality of algorithms for solving various MaOPs. The maximum number of iterations is set to 10000, and the maximum number of iterations would not exceed 13000. The population numbers of the algorithm were the same as in the settings of the original paper.

For the setting of some specific parameters in the algorithm, the performance was best when the number of direction vectors was set to 10 in DGEA. In DEA-GNG, the parameter α , used to determine the number of iterations for training GNG at each generation, was set to 0.1, and the default learning coefficient parameter *eps* was set to 0.314. In MaOEA-IGD, the number of evaluations for nadir point estimation DNPE was set to 100**N*. The remaining three algorithms did not involve specific parameters. These parameter settings were set empirically in the same manner as in the original studies. It was verified that the settings have achieved average performance over all the test suites in their studies.

5.3.3 Experimental results and analysis

Tables [16](#page-26-0) and [17](#page-28-0) show the results of different algorithms in terms of their average values of IGD and GD, where the values with better results are emphasized. When algorithm A was compared to algorithm B , "+" and "-" indicate the number of instances in which the results of *A* were significantly better and obviously worse than *B*, respectively. " $=$ " means that there was no statistically significant number of instances between the two algorithms. Therefore, as shown in the last row in the above tables, all comparison algorithms were compared to our proposed 1by1EA-F1. We can see from Table [16](#page-26-0) that 1by1EA-F1 achieved the 22 best IGD results among the 45 test instances, and the number of best results obtained by MaOEA-IGD, NSGA-II-SDR, DEA-GNG, DGEA, LMOCSO, and MSEA was 5, 0, 4, 4, 8, and 2, respectively. On LSMOP5/8/9, 1by1EA-F1 could perform optimally in almost all objective dimensions. On LSMOP1/3/6, our algorithm also achieved some excellent results. As shown in Table [17,](#page-28-0) the 31 best GD results among the 45 test instances were obtained by 1by1EA-F1, while the other algorithms only achieved 2, 10, 0, 0, 2, and 0. At the same time, 1by1EA-F1 fully demonstrated its superiority in the problem of objectives in all dimensions of LSMOP1/2/4/5/9. This means that 1by1EA-F1 had better performance on most test problems.

Another observation is that 1by1EA-F1 was more suitable for solving LSMOP5-9 where the hyperplane was not linear in terms of convergence and diversity of equilibrium solutions, while 1by1EA-F1 was suitable for almost all problems in terms of the convergence of the solution. One reason for this may be that for non-dominated solutions with nonlinear distributions, the algorithm can choose more randomly when selecting a single solution in a historical iteration. As the numbers of objective dimensions and decision variables increase, each single solution selected can exist more dispersed in the large-scale search space. This demonstrates the effectiveness of IFM in solving large-scale problems and in promoting population convergence. Another interesting point is that as LSMOP3 had the most complex multi-modal fitness landscape among the nine test problems, 1by1EA-F1 performed poorly on LSMOP3 with different objectives, which indicates that IFM has difficulty in always obtaining the best performance for large-scale test problems. This may be related to the shape of the PF. However, as this was an extreme case, after the above analysis, we can still conclude that 1by1EA-F1 can effectively solve large-scale MaOPs.

In the test functions for PF related to the convex landscape, the performance of MaOEA-IGD was better than the results of the test for the PF of linear. As MaOEA-IGD uses reference point-based assignment and the convex landscape has better reference point distribution, it was competitive on LSMOP5-8 with partial dimensionality. However, this also made it inferior to other algorithms on LSMOP1-4.

NSGA-II-SDR uses a niching technique based on Pareto dominance; so, it was very competitive in terms of convergence in comparison with other algorithms. However, for some specific LSMOPs, the inability to regulate the niche size affected its performance.

DEA-GNG did not seem to perform well in any LSMOPs except LSMOP2. As LSMOP2 is a partially separable problem, and PF has a mixed modality, GNG-based reference vectors in a complex problem can be adaptively adjusted; so, the population diversity is better. However, this may be because of the influence of the curvature information around the node. So, the algorithm performance could not be enhanced in large-scale problems.

While DGEA is also associated with large-scale problems, we found that it did not do well on almost all LSMOPs

Fig. 5 The mean of IGD and GD values on LSMOP1-9 with 15 objectives

other than LSMOP3. There is a pre-selection strategy in DGEA, which means that it is difficult to always obtain better performance using the same selection strategy on different testing problems. Furthermore, from the results of IGD and GD, DGEA could effectively maintain population diversity, but the convergence in large-scale space was still insufficient.

LMOCSO was proposed to solve large-scale MOPs, and it includes a novel particle update strategy that combines all the updated particles with the original particles. However, the results of the experiment suggest that this information retention strategy may not have as much impact as IFM.

MSEA replaces a solution in the population by improving the selection strategy. From the test results for LSMOPs, the limitations of MSEA were mainly related to the diversity of selection criteria. The normalization in 1by1EA improved the convergence of the algorithm, but the diversity of solution selection criteria may have made the algorithm transform a objective into the wrong scale when solving high-dimensional large-scale problems.

Figure [5](#page-30-0) shows the results of IGD and GD values calculated by these seven evolutionary algorithms in a clearer way when $m = 15$. Corresponding to the above table, when the objective number was 15, 1by1EA-F1 could obtain the lowest IGD value on LSMOP6/7/9/10. Then this algorithm could obtain the smaller GD value in almost all LSMOPs with different objectives. This indicates that 1by1EA-IFM can still adapt to most of the LSMOPs in a high-dimensional space, and performs excellently.

5.3.4 Statistical analysis of the results

Similar to the previous experiments, we still performed statistical analysis of the obtained experimental datasets in two steps. As shown in Table [18,](#page-31-0) the computed ranking through Friedman's test performed the difference among these algorithms. The results of this statistical test showed that 1by1EA-F1 obtained a significant difference with a *p*value tending to 0 infinitely under the significance level α = 0.05. The best ranking was obtained by 1by1EA-F1, followed by NSGA-II-SDR, DEAGNG, LMOCSO, MaOEA-IGD, MSEA, and DGEA respectively. This order was useful for the later Holm's test.

Next, we applied Holm's method to compare the best ranking algorithm with the remaing algorithms. Since the proposed 1byEA-F1 obtained the best ranking, it was considered as control algorithm. Table [19](#page-31-1) demonstrated the adjusted *p*-values by Holm post-hoc procedure, where the null hypothesis was rejected when the achieved *p*-value less than α / i . Analyzing these results, we concluded that all the *p*-values obtained by Holm post-hoc were lower than the adjusted α ($\alpha = \alpha/i$), which confirmed that the proposed

Table 18 Friedman test ranking for MaOEA-IGD, NSGA-II-SDR, DEA-GNG, DGEA, LMOCSO, MSEA, and 1by1EA-F1

Average Ranking
4.08
3.15
3.75
6.07
4.05
4.09
2.80

1by1EA-F1 statistically outperforms the other algorithms. Thus, the effectiveness of 1by1EA-IFM is verified.

Here, we use the above experimental results to explain again why the M-F1, M-F2 and M-R1 had the best performance in our proposed algorithms. First, in most evolutionary algorithms, the best individual in each environmental selection is retained. The introduction of IFM increases the uncertainty in environment selection. By retaining one, two, or three individuals from a fixed or random position in the historical iteration, the algorithm can avoid falling into the local optimum to a certain extent, thereby improving its convergence. Although the individual retained each time may not be the best, this selection method can indeed speed up the convergence. The above experimental results show that the GD values in most of the experiments were relatively low; so, 1by1EA-IFM made a contribution to the convergence.

The performance of 1by1EA-F1 was undoubtedly superior. This superiority comes from the fact that we only consider the use of information at a fixed location at the last generation, where the individuals are the closest to this generation. Thus, the value and inheritance of information were the highest. Moreover, selecting individuals at a fixed position in each iteration can increase the randomness of individual selection without making this randomness too messy; so, the 1by1EA-F1 had the best performance.

Table 19 Results obtained using Holm's post-hoc test

i	Algorithm	α/i	<i>p</i> -value of Holms	Null hypothesis
	NSGA-II-SDR	$0.05/1 = 0.0500$	0.025	Reject
2	DEAGNG	$0.05/2 = 0.0250$	0.024	Reject
3	LMOCSO	$0.05/3 = 0.0166$	0.012	Reject
4	MaOEA-IGD	$0.05/4 = 0.0125$	0.002	Reject
5	MSEA	$0.05/5 = 0.0100$	0.001	Reject
6	DGEA	$0.05/6 = 0.0083$	0.000	Reject

1by1EA-F1 used as control algorithm

6 Conclusion

In previous studies, few algorithms have been proposed that can reuse individual information in historical iteration populations. Inevitably, a lot of valuable information often appears in these algorithms. In this paper, a new 1by1EA was proposed based on this mechanism. It integrates information in historical populations and reuses them in subsequent iterations. Furthermore, this paper adopted the new algorithms specifically to solve the large-scale MaOPs. According to the six models of IFM, six algorithms were proposed. After comparing these six algorithms with the original 1by1EA, we found that 1by1EA-F1, 1by1EA-F2, and 1by1EA-R2 had the best overall performance. Among them, 1by1EA-F1 performed particularly well. Therefore, it was chosen for comparison in later experiments. By comparing the algorithms with the best algorithms in NSGA-III-IFM and MOEA/D-IFM on LSMOP1-9 and MOKP, we proved that 1by1EA-IFM could achieve the best performance in most experimental settings. Finally, we selected six relatively state-of-the-art evolutionary algorithms and verified them in the same large-scale MaOPs. The results verify the strong competitiveness of 1by1EA-F1.

In future research, we will focus on exploring why the performance of several other algorithms is not as good as that of 1by1EA-F1, and will make improvements to this problem on a series of MaOPs. Then, we will combine IFM with other MaOEAs to explore its ability to balance the diversity and convergence of algorithms. In addition to MOKP, for other problems such as TSP and scheduling, which have practical application value, we will consider incorporating the improved algorithms into them and observe whether the algorithms have the ability to resolve these problems. For the use of selection strategies, we will develop more strategies based on random selection and compare them with IFM to obtain better results. In addition, allowing the algorithm to learn adaptively to achieve dynamic optimization of the problem is also a direction worthy of consideration in the future.

Author Contributions Conceptualization: Yong Wang, Qian Zhang; methodology: Gai-Ge Wang; software: Gai-Ge Wang; validation: Qian Zhang; formal analysis: Yong Wang; investigation: Qian Zhang; resources: Gai-Ge Wang; data curation: Yong Wang; writing—original draft preparation: Qian Zhang; writing—review and editing: Gai-Ge Wang, Yong Wang; visualization: Yong Wang; supervision: Yong Wang; project administration: Gai-Ge Wang. All authors have read and agreed the final manuscript.

Declarations

Ethics approval and consent to participate This paper does not contain any study with human participants or animals performed by any author.

Consent for Publication All authors agreed with the content and that all gave explicit consent to submit.

Conflict of Interests All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

- 1. Zhang W, Hou W, Li C, Yang W, Gen M (2021) Multidirection update-based multiobjective particle swarm optimization for mixed no-idle flow-shop scheduling problem. Compl Syst Model Simul 1(3):176–197
- 2. Cai X, Hu Z, Chen J (2020) A many-objective optimization recommendation algorithm based on knowledge mining. Inf Sci 537:148–161
- 3. Lin Q, Liu S, Wong K-C, Gong M, Coello CAC, Chen J, Zhang J (2018) A clustering-based evolutionary algorithm for many-objective optimization problems. IEEE Trans Evol Comput 23(3):391–405
- 4. Hua Y, Liu Q, Hao K, Jin Y (2021) A survey of evolutionary algorithms for multi-objective optimization problems with irregular pareto fronts. IEEE/CAA J Autom Sin 8(2):303–318
- 5. Fan Q, Ersoy OK (2021) Zoning search with adaptive resource allocating method for balanced and imbalanced multimodal multiobjective optimization. IEEE/CAA J Autom Sin 8(6):1163–1176
- 6. Zhao F, Di S, Cao J, Tang J et al (2021) A novel cooperative multistage hyper-heuristic for combination optimization problems. Compl Syst Model Simul 1(2):91–108
- 7. Wang G-G, Deb S, Cui Z (2019) Monarch butterfly optimization. Neural Comput Applic 31(7):1995–2014
- 8. Tang J, Liu G, Pan Q (2021) A review on representative swarm intelligence algorithms for solving optimization problems: Applications and trends. IEEE/CAA J Autom Sin 8(10):1627– 1643
- 9. Houssein EH, Gad AG, Hussain K, Suganthan PN (2021) Major advances in particle swarm optimization: theory, analysis, and application, vol 63
- 10. Agarwal SK, Yadav S (2019) A comprehensive survey on artificial bee colony algorithm as a frontier in swarm intelligence. In: Ambient Communications and Computer Systems, Springer, pp 125–134
- 11. Roopa C, Harish B, Kumar SA (2019) Segmenting ecg and mri data using ant colony optimisation. Int J Artif Intell Soft Comput 7(1):46–58
- 12. Opara KR, Arabas J (2019) Differential evolution: a survey of theoretical analyses. Swarm Evol Comput 44:546–558
- 13. Zhang Q, Li H (2007) Moea/d: a multiobjective evolutionary algorithm based on decomposition. IEEE Trans Evol Comput 11(6):712–731
- 14. Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: Nsga-ii. IEEE Trans Evol Comput 6(2):182–197
- 15. Zitzler E, Laumanns M, Thiele L (2001) Spea2: improving the strength pareto evolutionary algorithm. TIK-report 103
- 16. Cui Z, Xue F, Cai X, Cao Y, Wang G-g, Chen J (2018) Detection of malicious code variants based on deep learning. IEEE Trans Ind Inform 14(7):3187–3196
- 17. Hu Z, Li Z, Dai C, Xu X, Xiong Z, Su Q (2020) Multiobjective grey prediction evolution algorithm for environmental/economic dispatch problem. IEEE Access 8:84162–84176
- 18. Cui Z, Zhao Y, Cao Y, Cai X, Zhang W, Chen J (2021) Malicious code detection under 5g hetnets based on a multi-objective rbm model. IEEE Netw 35(2):82–87
- 19. Shadkam E, Bijari M (2020) A novel improved cuckoo optimisation algorithm for engineering optimisation. Int J Artif Intell Soft Comput 7(2):164–177
- 20. Cai X, Cao Y, Ren Y, Cui Z, Zhang W (2021) Multi-objective evolutionary 3d face reconstruction based on improved encoder– decoder network. Inf Sci 581:233–248
- 21. Gao D, Wang G-G, Pedrycz W (2020) Solving fuzzy job-shop scheduling problem using de algorithm improved by a selection mechanism. IEEE Trans Fuzzy Syst 28(12):3265–3275
- 22. Wang G-G, Gao D, Pedrycz W (2022) Solving multi-objective fuzzy job-shop scheduling problem by a hybrid adaptive differential evolution algorithm. IEEE Trans Indust Inf
- 23. Han X, Han Y, Chen Q, Li J, Sang H, Liu Y, Pan Q, Nojima Y (2021) Distributed flow shop scheduling with sequence-dependent setup times using an improved iterated greedy algorithm. Compl Syst Model Simul 1(3):198–217
- 24. Wang G-G, Cai X, Cui Z, Min G, Chen J (2020) High performance computing for cyber physical social systems by using evolutionary multi-objective optimization algorithm. IEEE Trans Emerg Top Comput 8(1):20–30
- 25. Thakare AN, Bhagat L, Thomas A (2017) A self-organised routing algorithm for cognitive radio-based wireless sensor networks using biologically-inspired method. Int J Artif Intell Soft Comput 6(2):148–169
- 26. Dutta S, Das KN (2019) A survey on pareto-based eas to solve multi-objective optimization problems. In: Soft Computing for Problem Solving, Springer, pp 807–820
- 27. Liu Y, Gong D, Sun J, Jin Y (2017) A many-objective evolutionary algorithm using a one-by-one selection strategy. IEEE Trans Cybern 47(9):2689–2702
- 28. Wang G-G, Tan Y (2019) Improving metaheuristic algorithms with information feedback models. IEEE Trans Cybern 49(2):542–555
- 29. Wu T, An S, Han J, Shentu N (2022) An *ε*-domination based twoarchive 2 algorithm for many-objective optimization. J Syst Eng Electron 33(1):156–169
- 30. Prajapati A (2021) Two-archive fuzzy-pareto-dominance swarm optimization for many-objective software architecture reconstruction. Arab J Sci Eng 46(4):3503–3518
- 31. Chhabra JK (2018) Amarjeet: Fp-abc: Fuzzy-pareto dominance driven artificial bee colony algorithm for many-objective software module clustering. Comput Lang Syst Struct 51:1–21
- 32. Zheng W, Tan Y, Meng L, Zhang H (2018) An improved moea/d design for many-objective optimization problems. Appl Intell 48(10):3839–3861
- 33. Farias LR, Araújol AF (2019) Many-objective evolutionary algorithm based on decomposition with random and adaptive weights. In: 2019 IEEE international conference on systems, man and cybernetics (SMC), IEEE, pp 3746–3751
- 34. Sun Y, Yen GG, Yi Z (2018) Igd indicator-based evolutionary algorithm for many-objective optimization problems. IEEE Trans Evol Comput 23(2):173–187
- 35. Bader J, Zitzler E (2011) Hype: an algorithm for fast hypervolume-based many-objective optimization. Evol Comput 19(1):45–76
- 36. Li K, Wang R, Zhang T, Ishibuchi H (2018) Evolutionary manyobjective optimization: a comparative study of the state-of-the-art. IEEE Access 6:26194–26214
- 37. Deb K, Jain H (2013) An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: solving problems with box constraints. IEEE Trans Evol Comput 18(4):577–601
- 38. Liu R, Wang R, Bian R, Liu J, Jiao L (2021) A decompositionbased evolutionary algorithm with correlative selection mechanism for many-objective optimization. Evol Comput 29(2):269– 304
- 39. Palakonda V, Mallipeddi R, Suganthan PN (2021) An ensemble approach with external archive for multi-and many-objective optimization with adaptive mating mechanism and two-level environmental selection. Inf Sci 555:164–197
- 40. Zhao H, Zhang C, Zheng X, Zhang C, Zhang B (2022) A decomposition-based many-objective ant colony optimization algorithm with adaptive solution construction and selection approaches. Swarm Evol Comput 68:100977
- 41. Chen H, Cheng R, Wen J, Li H, Weng J (2020) Solving largescale many-objective optimization problems by covariance matrix adaptation evolution strategy with scalable small subpopulations. Inf Sci 509:457–469
- 42. He C, Cheng R, Yazdani D (2020) Adaptive offspring generation for evolutionary large-scale multiobjective optimization. IEEE Trans Syst Man Cybern Syst 52(2):
- 43. Tian Y, Zheng X, Zhang X, Jin Y (2019) Efficient largescale multiobjective optimization based on a competitive swarm optimizer. IEEE Trans Cybern 50(8):3696–3708
- 44. Gu Z-M, Wang G-G (2020) Improving nsga-iii algorithms with information feedback models for large-scale many-objective optimization. Futur Gener Comput Syst 107:49–69
- 45. Zhang Y, Wang G-G, Li K, Yeh W-C, Jian M, Dong J (2020) Enhancing moea/d with information feedback models for largescale many-objective optimization. Inf Sci 522:1–16
- 46. Saini N, Saha S (2021) Multi-objective optimization techniques: a survey of the state-of-the-art and applications. Eur Phys J Spec Top 230(10):2319–2335
- 47. Wei G (2017) Some cosine similarity measures for picture fuzzy sets and their applications to strategic decision making. Informatica 28(3):547–564
- 48. Jiang S, Yang S (2017) A strength pareto evolutionary algorithm based on reference direction for multiobjective and manyobjective optimization. IEEE Trans Evol Comput 21(3):329– 346
- 49. Yi J-H, Xing L-N, Wang G-G, Dong J, Vasilakos AV, Alavi AH, Wang L (2020) Behavior of crossover operators in nsga-iii for large-scale optimization problems. Inf Sci 509:470–487
- 50. Abed-alguni BH, Alawad NA, Barhoush M, Hammad R (2021) Exploratory cuckoo search for solving single-objective optimization problems. Soft Comput 25(15):10167–10180
- 51. Cheng R, Jin Y, Olhofer M et al (2016) Test problems for largescale multiobjective and many-objective optimization. IEEE Trans Cybern 47(12):4108–4121
- 52. Cai X, Xiao Y, Li M, Hu H, Ishibuchi H, Li X (2020) A gridbased inverted generational distance for multi/many-objective optimization. IEEE Trans Evol Comput 25(1):21–34
- 53. Liu Y, Wei J, Li X, Li M (2019) Generational distance indicatorbased evolutionary algorithm with an improved niching method for many-objective optimization problems. IEEE Access 7:63881– 63891
- 54. Chakkarapani K, Thangavelu T, Dharmalingam K, Thandavarayan P (2019) Multiobjective design optimization and analysis of magnetic flux distribution for slotless permanent magnet brushless dc motor using evolutionary algorithms. J Magn Magn Mater 476:524–537
- 55. Ishibuchi H, Imada R, Setoguchi Y, Nojima Y (2018) Reference point specification in inverted generational distance for triangular linear pareto front. IEEE Trans Evol Comput 22(6):961–975
- 56. Sun Y, Xue B, Zhang M, Yen GG (2018) A new two-stage evolutionary algorithm for many-objective optimization. IEEE Trans Evol Comput 23(5):748–761
- 57. Tian Y, Cheng R, Zhang X, Su Y, Jin Y (2018) A strengthened dominance relation considering convergence and diversity for evolutionary many-objective optimization. IEEE Trans Evol Comput 23(2):331–345
- 58. Jiao R, Zeng S, Li C, Ong Y-S (2021) Two-type weight adjustments in moea/d for highly constrained many-objective optimization. Inf Sci 578:592–614
- 59. Rostami S, Neri F (2017) A fast hypervolume driven selection mechanism for many-objective optimisation problems. Swarm Evol Comput 34:50–67
- 60. Panichella A (2019) An adaptive evolutionary algorithm based on non-euclidean geometry for many-objective optimization. In: Proceedings of the Genetic and Evolutionary Computation Conference, pp 595–603
- 61. Yuan J, Liu H-L, Gu F, Zhang Q, He Z (2020) Investigating the properties of indicators and an evolutionary many-objective algorithm using promising regions. IEEE Trans Evol Comput 25(1):75–86
- 62. Tian Y, Cheng R, Zhang X, Su Y, Jin Y (2019) A strengthened dominance relation considering convergence and diversity for evolutionary many-objective optimization. IEEE Trans Evol Comput 23:331–345
- 63. Liu Y, Ishibuchi H, Masuyama N, Nojima Y (2020) Adapting reference vectors and scalarizing functions by growing neural gas to handle irregular pareto fronts. IEEE Trans Evol Comput 24:439–453
- 64. Tian Y, Zheng X, Zhang X, Jin Y (2020) Efficient largescale multiobjective optimization based on a competitive swarm optimizer. IEEE Trans Cybern 50:3696–3708
- 65. Tian Y, He C, Cheng R, Zhang X (2021) A multistage evolutionary algorithm for better diversity preservation in multiobjective optimization. IEEE Trans Syst Man Cybern Syst 51:5880–5894

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Yong Wang received the Ph.D degree in computer science from Xi'an Jiaotong University, Xi'an, China, in 2009. He is currently an associate professor of software technology in the College of Computer Scicence and Technology, Ocean University of China, where he is also the Deputy Director of the College of Computer Scicence and Technology. He has authored or coauthored more than 30 referred papers in the area of artificial intelligence

and software engineering. His current research interests include intelligent computing and operations research.

Qian Zhang was born in November 1998 in Shandong Province, China. She successfully entered Qufu Normal University in 2016 and majored in computer science and technology. In June 2020, she successfully obtained a bachelor's degree from Qufu Normal University, and continued her graduate study at Ocean University of China. Her major is software engineering. Her main research direction is large-scale manyobjective optimization.

Gai-Ge Wang is an associate professor in Ocean University of China, China. His entire publications have been cited over 5200 times (Google Scholar). Thirteen and thirtynine papers are selected as Highly Cited Paper by Web of Science, and Scopus (till November 2019), respectively. One paper is selected as "Top Articles from Outstanding S&T Journals of China-F5000 Frontrunner". The latest Google h-index and i10 index are 40 and 78, respectively. He is senior member of SAISE, SCIEI, a member of

IEEE, IEEE CIS, ISMOST. He served as Editors-in-Chief of *OAJRC Computer and Communications*, Editorial Advisory Board Member of *Communications in Computational and Applied Mathematics (CCAM)*, Associate Editor of *IJCISIM*, an Editorial Board Member of *IEEE Access, Mathematics, IJBIC, Karbala International Journal of Modern Science,* and *Journal of Artificial Intelligence and Systems*. He served as Guest Editor for many journals including *Mathematics, IJBIC, FGCS, Memetic Computing* and *Operational Research*. His research interests are swarm intelligence, evolutionary computation, and big data optimization.