

基于多目标数据生成的昂贵多目标进化算法

黎建宇 詹志辉

(华南理工大学计算机科学与工程学院 广州 510006)

摘要 昂贵多目标优化问题是一类需要同时优化多个相互冲突且评估计算成本十分昂贵的目标的复杂优化问题,需要算法在计算资源受限的情况下尽可能找到目标值好且多样性好的一系列非支配解.进化计算方法是求解多目标优化问题的有效手段,但在求解昂贵多目标优化问题时仍面临多样性和收敛性这两个方面的挑战,即难以找到多样性好且收敛到全局最优的一系列解.针对上述挑战,本文提出了新型的基于多目标数据生成的昂贵多目标进化算法.本文的贡献点和创新点主要有以下三个方面.首先,本文提出并证明了非支配解生成定理,并基于此提出了多目标数据生成方法,以更有效地搜索到更多非支配解,提高算法的多样性.其次,本文提出了多种群多代理框架,使用多个代理模型替代评估成本昂贵的真实目标函数,并协同演化多个种群对多个代理模型进行协同求解,从而提高算法的收敛性.再次,基于上述提出的方法和框架,本文提出了基于多目标数据生成的昂贵多目标进化算法,以对昂贵多目标优化问题进行求解.为了验证算法性能,本文在两个著名测试集的共16个问题上进行了丰富的大量测试实验,并与现有的五个前沿算法进行对比.实验结果表明,本文提出的算法能在大部分问题上取得比所有对比算法都更好的性能,具有很好的有效性和高效性.

关键词 昂贵优化;多目标优化;进化计算;数据生成;协同演化

中图法分类号 TP18 DOI号 10.11897/SE.J.1016.2023.00896

Expensive Multi-Objective Evolutionary Algorithm with Multi-Objective Data Generation

LI Jian-Yu ZHAN Zhi-Hui

(School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006)

Abstract Expensive multi-objective optimization problem is a class of complex optimization problems that involve multiple objectives that are conflicted with each other and computationally expensive, which requires algorithms to find a set of non-dominated solutions as many and diversity as possible with limited computational resources. Although evolutionary computation algorithms are regarded as effective tools for solving multi-objective optimization problems, they still face the diversity and convergence challenges when solving expensive multi-objective optimization problems, i. e., difficulty to find a set of solutions that are of good diversity and converged to Pareto front. To address the diversity and convergence challenges, this paper proposes a novel multi-objective data generation-based expensive multi-objective evolutionary algorithm. The contributions and innovations of this paper can be mainly summarized in the following three aspects. First, this paper puts forward and proves the non-dominated solution generation theorem, and then proposes a multi-objective data generation method based on the theorem, so as to obtain more non-dominated solutions more efficiently for improving the solution diversity. Second, this paper proposes a multiple population for multiple surrogates framework that co-evolves multiple populations to

efficiently optimize multiple surrogates that are built for multiple real expensive objectives respectively. Third, based on the above proposed method and framework, this paper proposes the expensive multi-objective evolutionary algorithm based on multi-objective data generation for efficiently solving the expensive multi-objective optimization problem. To validate the algorithm performance, extensive experimental analyses are conducted on 16 problems from two well-known test sets in the related field with five existing state-of-the-art algorithms as competitors in this paper. The experimental results show that the algorithm proposed in this paper is able to achieve better metric values than all the compared algorithms on most of the problems, with good effectiveness and efficiency.

Keywords expensive optimization; multi-objective optimization; evolutionary computation; data generation; cooperative evolution

1 引言

昂贵多目标优化问题(Expensive Multi-objective Optimization Problem, EMOP)广泛存在于各种应用问题之中,是近年来多目标优化领域的前沿研究热点^[1-3]. 传统的多目标优化问题(Multi-objective Optimization Problem, MOP)是指一类需要同时优化多个相互冲突的目标的复杂优化问题^[4-6]. 在这类问题中,在一个目标上的性能提升会导致在另一个或多个目标上的性能下降,因此传统算法难以对这类问题进行有效求解. 进化计算(Evolutionary Computation, EC)^[7-8]是一类求解优化问题的高效方法,具有强大的问题求解能力. 因此,过去几十年来,研究者们借助进化计算方法提出和发展了众多求解 MOP 的多目标进化算法(Multi-Objective Evolutionary Algorithm, MOEA),并取得了不少的进展^[9-12]. 然而,相比于传统的 MOP, EMOP 不仅需要同时优化多个相互冲突的目标,而且各个目标函数的评估计算成本十分昂贵,导致可评估的次数十分有限. 因此, EMOP 要求算法在计算资源受限的情况下尽可能对多个相互冲突的目标进行求解. 比如,在翼型优化设计应用中,只计算 10 000 次真实的目标函数评估就需要约 188 天的时间(超过半年),这样的时间成本对研发和生产部门来说十分昂贵,难以接受^[13-14]. 可见, EMOP 作为一类计算资源受限的 MOP,比一般 MOP 更难以求解. 因此,如何对 EMOP 进行高效求解受到了各领域研究者的关注^[14-17].

由于 MOEA 在求解传统 MOP 问题上显示出

比其他方法更强的优势,很多研究者基于 MOEA 研究和设计新型算法对 EMOP 进行求解. 目前这些针对 EMOP 的研究主要集中在两个方面,分别是轻量级评估和高效算法设计^[3]. 在轻量级评估方面,计算复杂度更低、耗时更短的轻量级代理模型可用于替代昂贵的真实评估函数,从而减轻算法计算负担,提高算法的优化求解效率. 常见的代理模型类型包括高斯过程^[14-17]、径向基神经网络^[18-19]、神经网络^[20-23]、随机森林^[24]和支持向量机^[25]. 在高效算法设计方面,设计更高效的算法可以凭借更少的迭代搜索到满足需求的候选解,从而减少昂贵适应值的评估次数. 因此,研究者们研究和提出了各种算法框架、方法和策略,例如以基于分解的多目标优化框架为基础进行扩展的 MOEA/D-EGO^[17]、基于参考向量引导进化算法^[26](Reference Vector guided Evolutionary Algorithm, RVEA)框架的 KRVEA(Kriging-assisted RVEA)^[27]和 AB-MOEA(Adaptive Bayesian approach to MOEA)^[28].

尽管已有很多研究者研究和提出了上述求解 EMOP 的新型算法,这些算法在求解 EMOP 时仍存在多样性和收敛性这两个方面的挑战,即难以找到多样性好且收敛到帕累托(Pareto)前沿的一系列解. 针对这些困难和挑战,本文独辟蹊径,提出了新型的基于多目标数据生成的昂贵多目标进化算法(Expensive Multi-Objective Evolutionary Algorithm with Multi-objective Data Generation, EMOEA-MDG). EMOEA-MDG 算法主要基于本文提出的多目标数据生成(Multi-objective Data Generation, MDG)方法和多种种群多代理(Multiple Populations for Multiple Surrogates, MPMS)框架. 本文的研究

动机与思路主要有两方面. 首先, 近期的一些前沿工作表明数据生成和动态生长有助于求解昂贵的单目标优化问题^[18-21]. 受此启发, 本文创新性地提出面向多目标问题的数据生成方法(即 MDG), 从而生成多样化的候选个体, 提高算法求解 EMOP 时的搜索多样性. 其次, Zhan 等人^[29]提出的多种群多目标(Multiple Populations for Multiple Objectives, MPMO)框架在协同多种群求解传统 MOP 问题时能保持很好的搜索多样性和收敛性, 并被广泛应用到了众多应用问题之中, 例如作业车间调度^[30]、云工作流调度^[31]、供应链配置优化^[32]、疫情防控^[33]和航空机组排班^[34]. 因此, 本文基于 MPMO, 提出新型的 MPMS 框架, 通过协同多种群求解多个代理模型, 同时实现良好的搜索多样性和收敛性.

本文的创新点和贡献点可总结为以下三个方面:

(1) 本文提出并证明了非支配解生成定理, 用以分析多目标数据生成的可行性和有效性. 基于提出的非支配解生成定理, 本文提出了 MDG 方法, 从而生成更多样的非支配解数据, 提高算法的搜索多样性.

(2) 本文提出了新型的 MPMS 框架. MPMS 框架使用多个轻量级代理模型替代评估成本昂贵的真实目标函数, 并协同演化多个种群对多个代理模型进行协同求解, 实现良好的搜索多样性和收敛性.

(3) 基于上述提出的 MDG 和 MPMS, 本文提出了新型的 EMOEA-MDG, 以对 EMOP 进行求解. 此外, 本文还对 EMOEA-MDG 进行了时间复杂度的分析.

为了验证算法性能, 本文在 EMOP 领域两个著名测试集 DTLZ 和 WFG 的共 16 个问题上进行了丰富的大量实验测试, 并与现有的五个常用和前沿算法进行对比分析, 其中包括三个均发表在进化计算领域顶尖国际期刊《IEEE Transactions on Evolutionary Computation》的先进算法. 实验结果表明本文提出的 EMOEA-MDG 能在大部分问题上取得比所有对比算法都更好的性能, 具有很好的有效性和高效性.

本文第 2 节简单介绍 EMOP 及对相关工作进行简要综述; 第 3 节介绍本文提出的理论、MDG 方法、MPMS 框架和 EMOEA-MDG 算法; 第 4 节给出 EMOEA-MDG 算法与其他先进算法在 16 个问题上的实验对比, 并对算法运行效率和 MDG 的作用进行实验分析; 最后, 第 5 节对文本的工作进行总结和展望.

2 相关工作

2.1 昂贵多目标优化问题

EMOP 是指评估代价昂贵的一类 MOP. MOP 的相关概念定义同样适用于 EMOP. 以最小化的 MOP 为例, MOP 可被描述为

$$\text{Minimize } F(X) = \{f_1(X), f_2(X), \dots, f_M(X)\} \quad (1)$$

其中 X 是问题的决策空间 R^D 中的解, D 为问题的决策空间维度, f_1, f_2, \dots, f_M 是 M 个真实的目标函数, $F = [f_1, f_2, \dots, f_M] \in \Omega^M$, Ω^M 是问题的目标空间, M 是问题的目标空间维度. 基于式(1), MOP 有如下四个重要概念定义.

定义 1. 帕累托支配. 给定两个目标空间 Ω^M 中的向量 $U = [U_1, U_2, \dots, U_M]$ 和 $W = [W_1, W_2, \dots, W_M]$ ($U \neq W$). 如果对于任一整数 m ($1 \leq m \leq M$) 都有 $U_m \leq W_m$, 则称 U 支配(优于) W .

定义 2. 帕累托最优. 给定一个决策空间中的解 $X \in R^D$, 如果不存在 $X^* \in R^D$ 使得 $F(X^*)$ 支配 $F(X)$, 则称 X 为帕累托最优.

定义 3. 帕累托集合. 帕累托集合, 即 PS (Pareto Set), 定义为

$$PS = \{X \in R^D \text{ 且 } X \text{ 为帕累托最优}\} \quad (2)$$

定义 4. 帕累托前沿. 帕累托前沿, 即 PF (Pareto Front), 定义为

$$PF = \{F(X) \mid X \in PS\} \quad (3)$$

2.2 昂贵多目标优化的相关工作

现有针对昂贵多目标优化的研究主要集中在两个方面, 分别是轻量级评估和高效算法设计.

在轻量级评估方面, 计算代价更低的轻量级代理模型可用于替代昂贵的真实评估函数, 从而减轻算法计算负担, 提高算法的优化求解效率. 然而, 如果代理模型的预测不够精准, 与真实函数相差过大, 代理模型也会误导算法的搜索, 降低搜索精度. 因此, 很多研究者提出了建立更精准的代理模型的方法, 包括使用高斯过程^[14-17]、径向基神经网络^[18-19]、神经网络^[20-23]、随机森林^[24]和支持向量机^[25]. 然而, 由于多目标问题具有多个目标函数, 对多个目标函数进行建模的方式对算法的结果影响很大. 因此, 各种建立代理模型的方式也被广泛研究. 比如, Knowles^[35]提出在每次迭代随机采样一组权重将所有的目标函数聚合成一个目标函数, 然后对聚合后的函数进行代理建模, 用以评估个体的优劣. 而 Zhang 等人^[17]提出的 MOEA/D-EGO 算法则先将

多目标问题分解为若干个子问题,然后分别对每个子问题进行代理建模。Cai 等人^[36]和 Luo 等人^[37]也是先将多目标问题分解为若干个子问题,但在分解后会子问题分成若干组,对每一组内的多个子问题使用多任务的高斯过程进行协同代理建模。Chugh 等人^[27]提出的 K-RVEA 对每个目标函数进行分别建模,即建立的代理模型的数目等于目标函数的数目。此外,也有学者不对目标函数建模,而是选择其他方式建立代理模型。Guo 等人^[23]对多目标问题的常用评估指标进行建模预测,用预测的指标值来对个体优劣进行评估和区分。Pan 等人^[22]提出的 CSEA 则对候选解之间的优劣分类(即解之间的帕累托支配关系)进行建模,以辅助算法选择更优个体进行迭代演化。Yuan 等人^[38]则不仅对帕累托支配关系建立代理模型,还对 θ 支配关系建模,并协同两个模型辅助算法进行优化。

在高效算法设计方面,设计更高效的算法可以用更少的真实评估次数搜索到满足需求的候选解。因此,研究者们研究和提出了各种算法框架、方法和策略。Ponweiser 等人^[16]提出的 SMS-EGO (S-Metric Selection-based Efficient Global Optimization) 基于 EGO 框架进行改进,并结合了基于 S-Metric 和超体积的个体选择方法。MOEA/D-EGO 则基于经典的 MOEA/D 框架进行扩展,并提出基于 K-means 的选择方法,从而合理地选择个体进行昂贵评估和后续进化^[17]。而 K-RVEA 则基于近年提出的算法框架 RVEA,并提出基于不确定信息的模型管理方法,从而平衡算法的多样性和收敛性。Wang 等人^[28]提出的 AB-MOEA 同样由 RVEA^[26]框架扩展而成,但使用了新的自适应的个体采样和选择方法,从而实现更全局化的搜索。Guo 等人^[23]则基于现有基于指标的参考点自适应多目标进化算法^[39]并结合其提出的模型管理策略来对 EMOP 进行求解。除此之外,不同的选择方法、存档策略^[40-41]和局部搜索策略^[42]也被研究者们进行研究并用来提高算法效率。上述的工作从框架、选择方法、模型管理等多方面切入和研究面向 EMOP 的算法,而本文提出的 EMOEA-MDG 算法相比这些算法主要有两个新颖和重要的创新之处。首先,本文首次提出了使用数据生成的方法(即 MDG)来提高算法求解 EMOP 时的搜索多样性;其次,本文启发于针对传统多目标问题的 MPMS 框架,提出了新颖的 MPMS 框架,来提高算法求解 EMOP 时的多样性和收敛速度。因此,本文提出的

EMOEA-MDG 与现有工作有显著区别,而且数据生成的新思路 and MPMS 的新框架将有助于推动和启发相关领域的发展。

3 基于多目标数据生成的昂贵多目标进化算法

为了更好地对提出的算法进行介绍,本章首先给出非支配解生成定理,以便对数据生成的可行性及效果进行分析。然后,分别介绍提出的 MDG 方法以及 MPMS 框架。最后,结合上述的方法与框架,介绍提出的 EMOEA-MDG 算法。

3.1 多目标数据生成的理论分析

为了对数据生成的可行性进行理论分析,本文先给出非支配解生成定理,内容如下。

定理 1. 非支配解生成定理。给定对于 M 个目标函数 $F = [f_1, f_2, \dots, f_M]$ 的两个互不支配的解 U 和 W , 如果任一目标函数 f_i 的复合函数 $l_i(\alpha) = f_i(\alpha \times U + (1-\alpha) \times W)$ 都在区间 $\alpha \in [0, 1]$ 上单调, 其中 $1 \leq i \leq M$, 则基于 $\forall \lambda \in [0, 1]$ 生成的解 $Z = \lambda \times U + (1-\lambda) \times W$ 互不支配。

证明。不失一般性,这里假设下面讨论的所有目标函数都是最小化函数,即函数值越小越好。首先,由于 U 和 W 互不支配,则根据定义 1,至少存在两个目标函数(记为 f_i 和 f_j)使得 $f_i(U) > f_i(W)$ 且 $f_j(U) < f_j(W)$, 其中 $1 \leq i \leq M$ 且 $1 \leq j \leq M$ 且 $i \neq j$ 。对于 f_i 和 f_j 的复合目标函数 l_i 和 l_j , 我们有

$$\begin{cases} l_i(0) = f_i(0 \times U + (1-0) \times W) = f_i(W) \\ l_i(1) = f_i(1 \times U + (1-1) \times W) = f_i(U) \end{cases} \quad (4)$$

和

$$\begin{cases} l_j(0) = f_j(0 \times U + (1-0) \times W) = f_j(W) \\ l_j(1) = f_j(1 \times U + (1-1) \times W) = f_j(U) \end{cases} \quad (5)$$

由于 $f_i(U) > f_i(W)$, 我们有 $l_i(0) < l_i(1)$ 。而且,由于 l_i 在 $[0, 1]$ 上有单调性,则函数 l_i 在 $[0, 1]$ 上单调递增。同理,因为 $f_j(U) < f_j(W)$, 我们有 $l_j(0) > l_j(1)$ 和函数 l_j 在 $[0, 1]$ 上单调递减。任取 $0 \leq \lambda_1 < \lambda_2 \leq 1$, 记生成的解为 $Z_1 = \lambda_1 \times U + (1-\lambda_1) \times W$ 和 $Z_2 = \lambda_2 \times U + (1-\lambda_2) \times W$ 。由于 l_i 和 l_j 分别在 $[0, 1]$ 上单调递增和单调递减,则有

$$\begin{cases} l_i(0) \leq l_i(\lambda_1) < l_i(\lambda_2) \leq l_i(1) \\ l_j(0) \geq l_j(\lambda_1) > l_j(\lambda_2) \geq l_j(1) \end{cases} \quad (6)$$

即

$$\begin{cases} f_i(Z_1) = l_i(\lambda_1) < l_i(\lambda_2) = f_i(Z_2) \\ f_j(Z_1) = l_j(\lambda_1) > l_j(\lambda_2) = f_j(Z_2) \end{cases} \quad (7)$$

即 Z_1 在目标函数 f_i 上优于 Z_2 , 而 Z_2 在目标函数 f_j 上优于 Z_1 . 因此, 根据定义 1, 生成的解 Z_1 与 Z_2 互不支配, 即定理 1 得证. 证毕.

基于定理 1, 我们可以利用现有的非支配解生成更多的非支配解, 从而提高算法得到的非支配解集的多样性. 因此, 我们提出了多目标数据生成方法, 该方法将在下一部分进行介绍.

3.2 多目标数据生成方法

MDG 的主要思想是基于现有的非支配解进行线性组合生成更多样的非支配解. 为了使实际情况更符合定理 1 的前提条件, 即目标函数的复合函数能在两个非支配解组成的线性空间上有单调性, MDG 只选择最相近的两个互不支配的解为基础进行数据生成. MDG 利用如下公式生成新的解:

$$X_{\text{new}} = \frac{1}{2}(X_i + X_j) \quad (8)$$

其中, X_i 和 X_j 是两个不同的互不支配的解, X_{new} 是生成的新解. 为了进行更好的说明, 图 1 展示了多目标数据生成的效果. 由图 1 可见, 在目标函数 f_1 和 f_2 上, 解 X_i 和 X_j 都分别在一个函数上取得较小值而在另一个函数上取得较大值, 因此互不支配. 而基于解 X_i 和 X_j 生成的新解 X_{new} 在两个函数上都能取得较小值, 因此与 X_i 和 X_j 互不支配. 同时, 如图 1 右图所示, 在增加新解 X_{new} 后, 得到的非支配解集多样性更好, 且在帕累托前沿上的分布更加均匀.

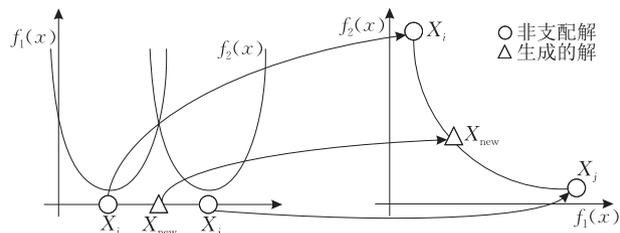


图 1 多目标数据生成的示意图

MDG 的伪代码如过程 1 所示. MDG 主要分为两步. 第一步是从现有种群 P 中获得非支配解集 (Non-dominated Solution Set, NSS), 如过程 1 的第 1 行, 其中 NSS 以符号 Ψ 表示. 然后, 过程 1 的第 2 行计算非支配解之间的欧式距离, 以便选出最近邻的另一个非支配解进行数据生成, 即对每一个非支配解, 如 X_i , 找到与它在决策空间中欧式距离最近的非支配解, 如 X_j ($X_i \neq X_j$), 并通过式 (8) 生成新数据, 加入到种群 P 中.

过程 1. 多目标数据生成方法.

1. 从现有种群 P 中选择出所有非支配解, 存到集合 Ψ ;
2. 计算 Ψ 中各个解之间在决策空间中的欧式距离;
3. FOR Ψ 中的每个解 X_i DO
4. 选择 Ψ 中与 X_i 欧式距离最近的个体, 记为 X_j ;
5. 基于 X_i 和 X_j 根据式 (8) 生成新个体;
6. 基于代理模型对新生成个体进行评估, 加入到 P ;
7. END FOR

3.3 多种群多代理框架

MPMS 框架的主要思想是首先使用多个代理模型分别替代真实昂贵的多个目标函数, 然后协同演化多个种群对多个代理模型进行协同求解.

多种群求解多代理的算法框架如图 2 所示. 对于有 M 个目标函数的昂贵多目标问题, 框架分别对 M 个目标函数建立代理模型, 并采用 M 个种群来协同求解这 M 个代理模型. 具体地, 该框架对问题的演化求解有三个主要步骤: (1) 每个种群内部进行交叉、变异等进化算子操作生成新个体; (2) 框架对 M 个种群的所有原有个体和新个体一并进行选择操作; (3) 用步骤 (2) 中选择得到的个体更新 M 个种群. 上述三个步骤将依次迭代执行, 直到算法停止.

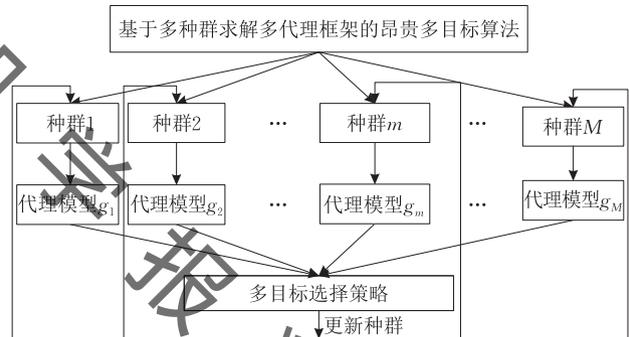


图 2 多种群求解多代理的昂贵多目标算法框架

对于框架中所需的选择操作, 本文提出了一种新型的两步选择策略. 该策略结合了现有文献中常用的非支配选择策略^[43]和基于拥挤距离的选择策略^[29]. 在本文提出的策略中, 第一步将对所有个体按非支配排序进行初步选择, 然后第二步按个体在目标空间的拥挤距离进行更精细的选择. 该策略的伪代码如过程 2 所示, 其中第 2 和第 1 行是对集合 P 和 Q 进行初始化. 第 3 到 6 行进行了迭代的非支配选择, 每次只从当前的 Q 中选择非支配的个体加入到目标集合 P , 并删除 Q 中被选中了的个体. 该过程一直迭代进行, 直到集合 P 中个体的数量大于等于预定的总种群规模 N (子种群的个体将在后续过程从 N 个个体中选择产生). 在第一步选择之后, 我们选择 P 中在目标空间具有更大拥挤距离的前 N 个个体作为新种群的个体, 并舍弃其它个体, 如

过程 2 的第 18 和 19 行所示. 过程 2 的第 9 到第 17 行计算每个个体的拥挤距离, 并用数组 d 记录结果. 其中, 每个目标函数的代理模型预测值最好和最差的个体的拥挤距离设定为最大 (MAX). 与传统 NSGA-II 等算法采用的选择策略^[43]相比, 本文提出的两步选择策略的主要不同点在于: 在第二步选择个体的时候, 前者只使用当前等级中的个体计算个体在目标空间中的拥挤距离并根据拥挤距离丢弃一些当前等级中较为“拥挤”的个体, 从而满足种群规模的限制, 而本文提出的方法使用整个种群的所有个体计算拥挤距离, 再根据拥挤距离丢弃较为“拥挤”的个体, 因此可以更全局化地提升种群分布在目标空间中的均匀程度.

过程 2. 两步选择策略.

1. 将所有种群的原有个体和新个体加入到 Q ;
2. 初始化 P 为空集;
3. WHILE P 中个体数量 $< N$ DO N 为最大种群规模
4. 将当前 Q 中所有非支配个体加入到 P ;
5. 删除当前 Q 中所有非支配个体;
6. END WHILE
7. 记 P 中个体数量为 P_N ;
8. 初始化 P_N 维的全零数组 d ;
//初始化个体密度值为 0
9. FOR $m=1$ to M // M 为代理模型的数目
10. 对 P 中个体按第 m 个代理模型的预测值排序;
11. 记排序索引为 L ; // $L(k)$ 表示排名 k 的个体索引
12. 设 $d_{L(1)}$ 和 $d_{L(P_N)}$ 为 MAX; // 拥挤距离为最大
13. $ms = |f_m(P_{L(P_N)}) - f_m(P_{L(1)})|$; // $|x|$
14. FOR $k=2$ to P_N-1 // M 为代理模型的数目
15. $d_{L(k)} = d_{L(k)} + |f_m(P_{L(k+1)}) - f_m(P_{L(k-1)})| / ms$;
16. END FOR
17. END FOR
18. 根据 d 值对 P_N 个个体进行降序排序;
19. 只保留 P 中具有更大 d 值的 N 个个体, 去掉其余个体;

此外, 框架的种群更新过程的伪代码如过程 3 所示. 由于框架采用 M 个子种群来求解对应 M 个目标函数的代理模型, 因此需要对 M 个子种群进行更新. 在更新第 m 个子种群 SP_m 的时候, 我们按照第 m 个代理模型的预测目标值, 从 P 中选择前 N_s 个个体复制作为新的第 m 个子种群 SP_m , 如过程 3 的第 3 和第 4 行所示. 其中, N_s 的计算过程如下:

$$N_s = \left\lfloor \frac{N}{M} \right\rfloor \quad (9)$$

其中 $\lfloor x \rfloor$ 表示小于 $x+1$ 的最大整数.

过程 3. 种群更新策略.

1. FOR $m=1$ to M
2. 清空对应第 m 个目标函数的子种群集合 SP_m ;
3. 对 P 中个体按第 m 个代理模型的预测值排序;

4. 将 P 中排名前 N_s 个个体复制到 SP_m ;
5. END FOR

3.4 EMOEA-MDG

EMOEA-MDG 算法基于上述 MDG 方法和 MPMS 框架, 整体流程如算法 1 所示. 算法的输入包括种群规模 N , 问题的目标数目 M , M 个真实的目标函数 f_1, f_2, \dots, f_M , 最大的真实评估次数 N_e . 初始化时真实评估的个体数目 NI , 优化过程中每两次真实评估的间隔代数为 $gmax$, 优化过程中每次真实评估的个体数目 u . 其中, 只有种群规模是算法参数, 其它输入均主要由问题及实际情况决定. 此外, 算法的输出是已进行真实昂贵评估的解的集合 A .

算法 1. EMOEA-MDG.

输入: 种群规模 N , 问题的目标数目 M , M 个真实的目标函数, 最大的真实评估次数 N_e , 初始化时真实评估的个体数目 NI , 优化过程中每两次真实评估的间隔代数 $gmax$, 优化过程中每次真实评估的个体数目 u

输出: 已评估的解集合 A .

1. 随机初始化 NI 个个体并进行真实评估, 作为初始的 A ;
2. $FEs = NI$; // 已消耗 NI 次昂贵的真实评估次数
3. 基于 A 中数据对 M 个目标分别构建 M 个代理模型;
4. 随机初始化 N 个个体, 作为初始的 P ;
5. 基于 M 个代理模型评估 P 中每个个体的 M 个目标值;
6. 基于式(9)计算 N_s ; // N_s 为子种群规模
7. 初始化 M 个子种群, 即 SP_1 到 SP_M ; // 过程 3
8. $g = 1$;
9. WHILE (A 中个体数量 $< N_e$) DO
// N_e 为最大的真实评估次数
10. FOR $m=1$ to M
11. 基于进化算子从 SP_m 生成 N_s 个个体;
12. 基于代理模型评估生成个体的 M 个目标值;
13. 将生成的 N_s 个个体加入到 P ;
14. END FOR
15. 对 P 进行数据生成, 将生成个体加入 P ; // 过程 1
16. 对 P 进行多目标选择; // 过程 2
17. 基于选择后的 P 更新 M 个子种群; // 过程 3
18. IF ($g = gmax$) THEN
19. 从 P 中选出非支配解;
20. 从非支配解中随机选择 u 个解进行真实评估;
21. $FEs = FEs + u$;
22. 将评估后的解加入到 A 中;
23. 基于 A 中的数据更新 M 个代理模型;
24. $g = 1$;
25. ELSE
26. $g = g + 1$;
27. END IF
28. END WHILE
29. 输出 A ;

算法 1 主要包括两个部分:初始化部分和演化求解部分,分别为算法 1 的第 1 到 8 行和第 9 到 28 行.初始化部分主要包括评估数据的初始化(第 1 和第 2 行),代理模型的初始化(第 3 行),以及种群和子种群的初始化(第 4 到第 7 行).在演化求解部分,算法主要包括子种群演化部分(第 10 到 14 行),数据生成部分(第 15 行),个体选择部分(第 16 行),子种群更新部分(第 17 行),以及真实评估部分(第 18 到 27 行).其中,子种群演化部分使用包括交叉和变异的进化算子生成新个体,见第 11 行.而在真实评估部分,由于真实评估十分昂贵,算法只有有限的真实评估次数.因此算法每隔 $gmax$ 代进行一次真实评估,每次评估 u 个候选解.被进行真实评估的解均通过在种群 P 中的非支配解中随机选择产生.新增的真实评估的解会加入到 A 中,且更新后的 A 会被用于更新所有的 M 个代理模型,以提高模型的预测准确率,更好地辅助种群演化求解过程.上述的这些过程,将迭代循环直至真实评估次数达到最大次数 N_e .最后,算法输出已评估的解集 A .

3.5 时间复杂度分析

为了简便,在对算法进行时间复杂度分析前,我们先对过程 1 到过程 3 进行分析.另外,由于代理模型的建模和预测时间复杂度独立于算法,且与选择的模型类型相关,为了不失一般性,本文对算法的时间复杂度分析不考虑代理模型.

过程 1 的时间复杂度分析如下.由于过程 1 的时间复杂度与 Ψ 的规模有关,而 Ψ 的规模不固定,我们以最坏情况(即 Ψ 与 P 等价)为例分析过程 1 的时间复杂度.在第 1 行,从 P 中选择非支配解需要两两比较所有解的 M 个目标值,且 P 的规模为 N ,因此该行时间复杂度为 $O(N^2 M)$.同时,在最坏情况下,第 2 行的时间复杂度为 $O(N^2 D)$,其中 D 为问题维度.第 4 行是选择欧式距离最小的个体,因此时间复杂度只需 $O(\log N)$,第 5 行是单个个体生成,时间复杂度为 $O(D)$.由于第 4 和第 5 行会进行循环,循环次数与 Ψ 规模有关,当最坏情况下(即 Ψ 与 P 等价) Ψ 会有 N 个解,第 4 和第 5 行需循环 N 次,对应时间复杂度为 $O(N \log N + ND)$.综上,过程 1 的时间复杂度为 $O(N^2 M + N^2 D + N \log N + ND)$,合并简化后为 $O(N^2 M + N^2 D)$.

过程 2 的时间复杂度主要取决于第 15 行双层循环的拥挤距离计算.由于有嵌套循环,第 15 行的时间复杂度为 $O(M \times P_N)$,其中 P_N 是 P 中个体数目.由于 P 是 Q 的子集, P 中个体数目不会大于 Q 中个体数目,即 $P_N < 2N$,因此,过程 2 的时

间复杂度可简化为 $O(MN)$.

过程 3 的时间复杂度主要取决于第 3 行的排序.由于该排序会循环 M 次,第 3 行的时间复杂度为 $O(MN \log N)$,即过程 3 的时间复杂度为 $O(MN \log N)$.

算法 1 的时间复杂度主要取决于主循环的第 10 到第 17 行.由于第 10 到第 14 行主要是 M 个种群的个体生成,而每个种群产生 N_s 个个体,因此第 10 到第 14 的时间复杂度为 $O(MN_s D)$, D 为每个个体的维度.根据式(9), $MN_s \approx N$,因此,该时间复杂度可记为 $O(ND)$.第 15 到 17 行主要为过程 1、过程 2 和过程 3 这三个过程的执行,因此每次执行第 15 到 17 行的总时间复杂度为 $O(N^2 M + N^2 D + MN + MN \log N)$,简化后为 $O(N^2 M + N^2 D)$.记第 10 到第 17 行所需的循环次数为 $T(T = (N_e - NI) / u)$,算法中第 10 到第 17 行的时间复杂度为 $O(TN^2 M + TN^2 D)$,即算法 1 的时间复杂度为 $O(TN^2 M + TN^2 D)$.

4 实验分析

4.1 实验设置

本文采用现有文献常用的两个著名多目标问题测试集来测试本文提出的 EMOEA-MDG 算法的性能.其中,本文采用来自 DTLZ 测试集^[44]的 7 个测试问题(DTLZ1 到 DTLZ7)和来自 WFG 测试集^[45]的 9 个测试问题(WFG1 到 WFG9),共 16 个测试问题.按照现有的文献[27-28],测试问题的相关设置如下:在实验中,每个问题的目标数均设置为 3,即 $M=3$,问题规模均为 10 维,即 $D=10$.每个问题的最大真实评估次数 N_e 为 230,包括 110 次($11 \times D$)在初始化之前的评估次数,即 $NI=110$.因此,每个算法在优化过程中最多只能再选择 $230 - 110 = 120$ 个候选个体进行真实评估.上述如此有限的真实评估次数能有助于考察算法求解昂贵多目标优化问题的高效性.同时,为了更好地考察算法的优越性,本文采用了 5 个现有文献中常用和前沿的算法进行比较.这 5 个算法分别是 SMS-EGO^[16]、MOEA/D-EGO^[17]、CSEA^[22]、K-RVEA^[27] 和 AB-MOEA^[28].其中,MOEA/D-EGO、CSEA 和 K-RVEA 均发表在进化计算领域顶尖国际期刊《IEEE Transactions on Evolutionary Computation》.每个算法都独立运行 30 次,然后将得到的结果进行比较.在实验中,每个算法的种群规模均设置为 100,即 $N=100$.因此,根据式(9),本文提出的算法的子种群规模为 34.所有算法每次最多选择 $u=5$ 个个体进行真实评估,同时

每两次的真实评估之间需要间隔 $gmax=20$ 代的代理模型辅助的优化过程. 为保持实验结果的可靠性和准确性, 所有的实验均在公开的多目标优化平台 PlatEMO^[46] 上进行.

4.2 评价指标

本文采用两个常用的多目标优化评估指标对算法性能进行评估和比较, 分别是 IGD (Inverted Generalized Distance)^[5] 和 HV (Hypervolume)^[5]. 指标 IGD 的计算公式如下:

$$IGD(S, S^*) = \frac{\sum_{x \in S^*} \min_{y \in S} dist(x, y)}{|S^*|} \quad (10)$$

其中, S 是算法得到的已被真实评估的解的集合, S^* 是在真实帕累托前沿上均匀采样得到的参考点的集合, $|S^*|$ 是表示集合 S^* 包含的元素的数量, $dist(x, y)$ 表示解 x 与解 y 在目标空间的欧式距离.

IGD 衡量从理想解集 S^* 到算法给出的解集的总体距离. 如果解集 S 中的解无法很好地收敛靠近

理想解集 S^* , 则集合 S 和 S^* 中的解之间的总体距离会很大, 导致很大的 IGD 值. 如果 S 中的解只收敛靠近 S^* 中部分的解, 则集合 S 和 S^* 总体中的解之间的总体距离仍会很大, 导致很大的 IGD 值. 因此, IGD 指标可以很好地综合反映出算法得到的解集的收敛性和多样性, 且 IGD 值越小越好. 指标 HV 衡量由参考点与算法得到的解在目标空间中围成的区域的体积. 越大的 HV 表示算法找到的帕累托前沿的收敛程度和范围越好. 在实验中, 本文采用每个算法独立运行 30 次后的 30 个 IGD 和 HV 值的均值和方差进行比较. 同时, 本文采用显著性水平设置为 0.05 的 Wilcoxon 秩和检验来对算法的结果进行对比检验. 根据 Wilcoxon 秩和检验, 符号“+”、“≈”和“-”会被分别用于表示本文提出算法比对比算法具有显著好、无显著差异和显著差的结果.

4.3 与前沿算法的实验结果对比

表 1 和表 2 为 EMOEA-MDG 与五个对比算法在 16 个问题上的结果. 其中, 加粗的指标数据表示

表 1 EMOEA-MDG 与现有前沿算法求解 DTLZ 与 WFG 中共 16 个多目标问题的 IGD 指标均值与标准差结果

测试问题	EMOEA-MDG	SMS-EGO	MOEA/D-EGO	CSEA	K-RVEA	AB-MOEA
DTLZ1	1.1869E+02 (2.41E+01)	1.1347E+02 (2.01E+01)=	8.6400E+01 (1.80E+01)-	6.4687E+01 (1.57E+01)-	8.9025E+01 (1.71E+01)-	8.7450E+01 (2.15E+01)-
DTLZ2	1.2983E-01 (2.20E-02)	3.0720E-01 (3.54E-02)+	3.0245E-01 (3.20E-02)+	2.2758E-01 (2.88E-02)+	1.4795E-01 (3.17E-02)+	1.6526E-01 (3.27E-02)+
DTLZ3	3.0240E+02 (6.22E+01)	2.5775E+02 (7.14E+01)-	2.0020E+02 (2.45E+01)-	1.8057E+02 (6.09E+01)-	2.5657E+02 (5.64E+01)-	2.6932E+02 (5.66E+01)-
DTLZ4	5.7063E-01 (1.36E-01)	7.3493E-01 (1.28E-01)+	6.2988E-01 (5.99E-02)=	4.9233E-01 (1.75E-01)-	3.9165E-01 (1.31E-01)-	4.1225E-01 (1.03E-01)-
DTLZ5	2.7070E-02 (6.43E-03)	1.9911E-01 (4.15E-02)+	2.5429E-01 (2.82E-02)+	1.3712E-01 (3.57E-02)+	1.0613E-01 (3.04E-02)+	1.1767E-01 (3.49E-02)+
DTLZ6	5.4236E+00 (3.02E-01)	5.6097E+00 (4.28E-01)+	2.0928E+00 (6.98E-01)-	5.2196E+00 (6.79E-01)=	3.4161E+00 (4.33E-01)-	3.7222E+00 (5.57E-01)-
DTLZ7	2.7389E-01 (1.46E-01)	2.8672E+00 (1.11E+00)+	2.4545E-01 (8.03E-02)=	1.9225E+00 (7.79E-01)+	1.6903E-01 (5.27E-02)-	5.0307E-01 (2.74E-01)+
WFG1	2.0744E+00 (1.07E-01)	2.2727E+00 (9.41E-02)+	2.2033E+00 (8.73E-02)+	1.7860E+00 (1.02E-01)-	1.8347E+00 (1.23E-01)-	1.8289E+00 (1.04E-01)-
WFG2	4.5053E-01 (5.30E-02)	6.1088E-01 (5.71E-02)+	6.9935E-01 (4.76E-02)+	5.6725E-01 (7.59E-02)+	4.4770E-01 (6.84E-02)=	4.9895E-01 (6.61E-02)+
WFG3	3.0270E-01 (3.97E-02)	5.9617E-01 (6.02E-02)+	6.2612E-01 (3.42E-02)+	5.1650E-01 (6.72E-02)+	4.7863E-01 (6.42E-02)+	4.3471E-01 (5.71E-02)+
WFG4	4.6864E-01 (6.41E-02)	5.4665E-01 (1.92E-02)+	5.9271E-01 (3.89E-02)+	4.6519E-01 (4.44E-02)=	4.8584E-01 (3.88E-02)+	4.7278E-01 (2.26E-02)=
WFG5	4.6733E-01 (9.31E-02)	5.9776E-01 (3.83E-02)+	5.9081E-01 (5.35E-02)+	5.6420E-01 (3.92E-02)+	4.5011E-01 (5.16E-02)=	4.8749E-01 (4.80E-02)=
WFG6	5.8548E-01 (7.83E-02)	7.8972E-01 (2.59E-02)+	8.1488E-01 (3.73E-02)+	7.2202E-01 (4.12E-02)+	7.4342E-01 (5.55E-02)+	7.5936E-01 (3.97E-02)+
WFG7	5.4816E-01 (2.85E-02)	6.9362E-01 (3.29E-02)+	6.7580E-01 (2.70E-02)+	5.9390E-01 (4.80E-02)+	6.2683E-01 (4.75E-02)+	5.9891E-01 (2.68E-02)+
WFG8	5.8998E-01 (4.03E-02)	7.6671E-01 (2.77E-02)+	8.6678E-01 (3.89E-02)+	7.4591E-01 (5.51E-02)+	7.5626E-01 (4.42E-02)+	7.0455E-01 (4.84E-02)+
WFG9	6.7954E-01 (8.17E-02)	8.6535E-01 (4.15E-02)+	8.1503E-01 (6.42E-02)+	6.9810E-01 (8.42E-02)=	7.1307E-01 (8.48E-02)=	6.9742E-01 (7.19E-02)=
+/-/=	NA	14/1/1	11/2/3	9/3/4	7/3/6	8/3/5
最优结果占比/%	43.75	0	6.25	25	25	0

表 2 EMOEA-MDG 与现有前沿算法求解 DTLZ 与 WFG 中共 16 个测试问题的 HV 指标均值与标准差结果

测试问题	EMOEA-MDG	SMS-EGO	MOEA/D-EGO	CSEA	K-RVEA	AB-MOEA
DTLZ1	0.0000E+00 (0.00E+00)	0.0000E+00 (0.00E+00)=				
DTLZ2	4.5430E-01 (2.74E-02)	1.6228E-01 (5.69E-02)+	1.4729E-01 (4.75E-02)+	2.8276E-01 (5.76E-02)+	3.9914E-01 (5.24E-02)+	3.5424E-01 (6.28E-02)+
DTLZ3	0.0000E+00 (0.00E+00)	0.0000E+00 (0.00E+00)=				
DTLZ4	8.0225E-02 (8.05E-02)	1.8603E-02 (2.62E-02)+	1.0369E-02 (1.57E-02)+	1.5646E-01 (7.71E-02)-	1.2421E-01 (1.19E-01)=	8.2269E-02 (8.00E-02)=
DTLZ5	1.8323E-01 (5.45E-03)	4.7749E-02 (2.24E-02)+	1.7211E-02 (1.37E-02)+	7.2227E-02 (3.19E-02)+	1.0271E-01 (2.86E-02)+	8.9282E-02 (3.44E-02)+
DTLZ6	0.0000E+00 (0.00E+00)	0.0000E+00 (0.00E+00)=				
DTLZ7	1.8386E-01 (1.51E-02)	4.1168E-04 (9.12E-04)+	2.0102E-01 (2.29E-02)-	3.4549E-02 (4.34E-02)+	2.4251E-01 (8.22E-03)-	2.0733E-01 (2.29E-02)-
WFG1	1.7741E-02 (2.32E-02)	1.0833E-03 (4.51E-03)+	6.4020E-03 (1.31E-02)+	1.3334E-01 (4.99E-02)-	1.3104E-01 (3.86E-02)-	1.1684E-01 (5.16E-02)-
WFG2	7.7970E-01 (2.39E-02)	6.7076E-01 (2.94E-02)+	6.3657E-01 (1.56E-02)+	6.9402E-01 (3.95E-02)+	7.3368E-01 (3.73E-02)+	7.0399E-01 (3.16E-02)+
WFG3	2.7407E-01 (2.17E-02)	1.7128E-01 (2.33E-02)+	1.5157E-01 (1.55E-02)+	1.9002E-01 (2.23E-02)+	2.1124E-01 (2.49E-02)+	2.1607E-01 (2.61E-02)+
WFG4	3.6260E-01 (2.31E-02)	3.1792E-01 (7.28E-03)+	3.2761E-01 (1.27E-02)+	3.8361E-01 (1.93E-02)-	3.5320E-01 (2.00E-02)=	3.6555E-01 (1.37E-02)=
WFG5	3.6714E-01 (5.29E-02)	2.9080E-01 (2.23E-02)+	3.5446E-01 (1.99E-02)=	3.2857E-01 (2.33E-02)+	3.8175E-01 (2.51E-02)=	3.5043E-01 (2.83E-02)=
WFG6	3.0422E-01 (4.59E-02)	2.0799E-01 (1.09E-02)+	2.4076E-01 (1.87E-02)+	2.5546E-01 (3.23E-02)+	2.4172E-01 (2.65E-02)+	2.2779E-01 (1.24E-02)+
WFG7	3.2362E-01 (1.44E-02)	2.5256E-01 (8.92E-03)+	2.7811E-01 (1.18E-02)+	3.2437E-01 (2.69E-02)=	2.8506E-01 (2.45E-02)+	3.1024E-01 (1.73E-02)+
WFG8	3.0517E-01 (1.92E-02)	2.3345E-01 (1.39E-02)+	1.9874E-01 (1.32E-02)+	2.5168E-01 (2.41E-02)+	2.5998E-01 (1.80E-02)+	2.6188E-01 (1.72E-02)+
WFG9	2.5949E-01 (3.58E-02)	1.9628E-01 (1.93E-02)+	0.0000E+00 (0.00E+00)+	0.0000E+00 (0.00E+00)+	2.4688E-01 (3.94E-02)=	2.5844E-01 (2.94E-02)=
+/-/+	NA	13/3/0	11/4/1	9/4/3	7/7/2	7/7/2
最优结果占比/%	62.5	18.75	18.75	43.75	31.25	18.75

对应算法获得的指标结果在所有算法中最好。表 1 和表 2 的实验数据表明,本文提出的 EMOEA-MDG 在总体上显著优于所对比的五个算法,由表 1 可知,EMOEA-MDG 分别在 43.75% 和 62.5% 的问题上取得最优的 IGD 值和 HV 值,远远高于其它对比算法。同时,根据 IGD 数据结果的 Wilcoxon 秩和检验,EMOEA-MDG 分别在 14、11、9、7 和 8 个问题上显著优于 SMS-EGO、MOEA/D-EGO、CSEA、K-RVEA 和 AB-MOEA。根据 HV 数据结果的 Wilcoxon 秩和检验,EMOEA-MDG 分别在 13、11、9、7 和 7 个问题上显著优于 SMS-EGO、MOEA/D-EGO、CSEA、K-RVEA 和 AB-MOEA。上述的结果表明了 EMOEA-MDG 优越的总体性能。此外,实验数据也表明,相比于 DTLZ 问题,EMOEA-MDG 在求解 WFG 问题的优越性更加明显。在 WFG 的 9 个问题中,EMOEA-MDG 只在 3 个问题上未取得最优的 IGD 值,即问题 WFG1、WFG4 和 WFG5。然而,根

据 Wilcoxon 秩和检验,EMOEA-MDG 在 WFG4 和 WFG5 上的 IGD 结果与最优的结果(即分别为 CSEA 在 WFG4 的结果和 K-RVEA 在 WFG5 的结果)并无显著性差异。可见,EMOEA-MDG 在解决 WFG 测试集问题时具有很强的优越性。而且,DTLZ 测试集与 WFG 测试集的差别主要在于 DTLZ 测试集的问题变量可分,而 WFG 测试集的问题变量不可分,这会导致对 WFG 测试集问题的优化求解更加困难。因此,实验结果反映出尽管现有的前沿算法在 DTLZ 问题上取得很好的结果,在求解更复杂的 WFG 测试集问题时仍然面临着较大挑战。而本文提出的 EMOEA-MDG 即使在更复杂的 WFG 测试集的问题中仍能保持良好的性能,得到比现有算法更好的结果。综上所述,本文提出的 EMOEA-MDG 具有优越的性能。

4.4 算法的运行效率分析

为了对算法的运行效率进行考察,本文记录了

不同算法分别在求解 16 个问题时的 30 次独立运行的平均耗时,结果如图 3 所示. 由图 3 可知,EMOEA-MDG 所需耗时明显比 SMS-EGO、MOEA/D-EGO 和 CSEA 都更短,与 K-RVEA 和 AB-MOEA 的所需耗时有着较小的差距. 由此可见,本文提出的 EMOEA-MDG 的运行求解效率在所有六个算法中排名前三. 而且,考虑到 EMOEA-MDG 比起 K-RVEA 和

AB-MOEA 有更好的优化结果,EMOEA-MDG 相比 K-RVEA 和 AB-MOEA 所需的额外耗时可以接受. 此外,不同于 SMS-EGO 和 CSEA 等耗时情况会随问题变化而变化的算法,本文提出的 EMOEA-MDG 在求解不同的问题时所耗的耗时差别不大,具有较好的稳定性. 因此,综上所述,本文提出的 EMOEA-MDG 具有较好且稳定的问题求解和运行效率.

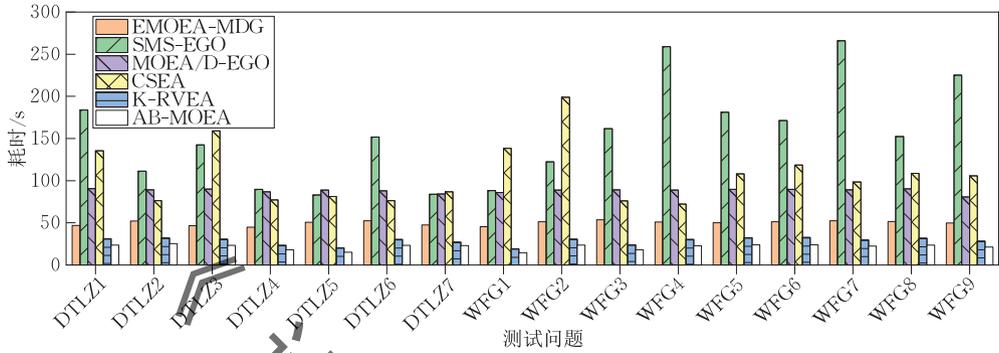


图 3 不同算法分别在求解 16 个问题时的 30 次独立运行的平均耗时结果

4.5 数据生成的作用分析

为了对提出的数据生成方法进行实验分析,本文设计了不使用 MDG 的算法变种,记为 EMOEA-w/o-MDG,并将其与原来的 EMOEA-MDG 进行比较. EMOEA-MDG 与 MDG 的 *IGD* 和 *HV* 结果如表 3 所示. 表 3 的结果显示,尽管 EMOEA-MDG 和 EMOEA-w/o-MDG 能在差不多数量的问题上取得最优的结果,EMOEA-MDG 还是在总体上显著优于 EMOEA-

w/o-MDG. 根据 Wilcoxon 秩和检验,EMOEA-MDG 分别在 6 个和 7 问题上取得显著优于 EMOEA-w/o-MDG 的 *IGD* 和 *HV* 结果,并分别在另外 7 和 6 个问题上与 EMOEA-w/o-MDG 结果无显著性差别. 由此可见,使用 MDG 能在 6 到 7 个问题上显著提高算法的性能,同时在其它不同问题上仍能保持算法的性能. 因此,MDG 具有较好的通用性和有效性,有助于算法对昂贵多目标问题进行高效求解.

表 3 EMOEA-MDG 与 EMOEA-w/o-MDG 在 16 个测试问题上的结果比较

测试问题	IGD 指标结果		测试问题	HV 指标结果	
	EMOEA-MDG	EMOEA-w/o-MDG		EMOEA-MDG	EMOEA-w/o-MDG
DTLZ1	1.1869E+02(2.41E+01)	1.1376E+02(2.38E+01) =	DTLZ1	0.0000E+00(0.00E+00)	0.0000E+00(0.00E+00) =
DTLZ2	1.2983E-01(2.20E-02)	1.8761E-01(2.22E-02) +	DTLZ2	4.5430E-01(2.74E-02)	3.9681E-01(3.28E-02) +
DTLZ3	3.0240E+02(6.22E+01)	2.9623E+02(6.86E+01) =	DTLZ3	0.0000E+00(0.00E+00)	0.0000E+00(0.00E+00) =
DTLZ4	5.7063E-01(1.36E-01)	5.5769E-01(1.10E-01) =	DTLZ4	8.0225E-02(8.05E-02)	4.9566E-02(6.42E-02) +
DTLZ5	2.7070E-02(6.43E-03)	1.8363E-01(9.01E-02) +	DTLZ5	1.8323E-01(5.45E-03)	8.3976E-02(5.86E-02) +
DTLZ6	5.4236E+00(3.02E-01)	3.8502E+00(4.80E-01) -	DTLZ6	0.0000E+00(0.00E+00)	0.0000E+00(0.00E+00) =
DTLZ7	2.7389E-01(1.46E-01)	1.9961E-01(7.48E-02) -	DTLZ7	1.8386E-01(1.51E-02)	2.0451E-01(2.40E-02) -
WFG1	2.0744E+00(1.07E-01)	1.9146E+00(1.34E-01) -	WFG1	1.7741E-02(2.32E-02)	9.5812E-02(4.78E-02) -
WFG2	4.5053E-01(5.30E-02)	4.5703E-01(6.32E-02) =	WFG2	7.7970E-01(2.39E-02)	7.8154E-01(3.35E-02) =
WFG3	3.0270E-01(3.97E-02)	3.7807E-01(5.09E-02) +	WFG3	2.7407E-01(2.17E-02)	2.5799E-01(2.30E-02) +
WFG4	4.6864E-01(6.41E-02)	4.8154E-01(2.73E-02) =	WFG4	3.6260E-01(2.31E-02)	3.6091E-01(1.73E-02) =
WFG5	4.6733E-01(9.31E-02)	4.2533E-01(8.59E-02) =	WFG5	3.6714E-01(5.29E-02)	4.0550E-01(4.59E-02) -
WFG6	5.8548E-01(7.83E-02)	7.1789E-01(6.41E-02) +	WFG6	3.0422E-01(4.59E-02)	2.4979E-01(3.16E-02) +
WFG7	5.4816E-01(2.85E-02)	6.1172E-01(3.12E-02) +	WFG7	3.2362E-01(1.44E-02)	3.1252E-01(1.79E-02) +
WFG8	5.8998E-01(4.03E-02)	6.4328E-01(4.09E-02) +	WFG8	3.0517E-01(1.92E-02)	2.7798E-01(1.72E-02) +
WFG9	6.7954E-01(8.17E-02)	6.8163E-01(6.63E-02) =	WFG9	2.5949E-01(3.58E-02)	2.6352E-01(3.36E-02) =
+ / = / -	NA	6 / 7 / 3	+ / = / -	NA	7 / 6 / 3
最优结果占比 / %	56.25	43.75	最优结果占比 / %	50	50

5 结 论

本文提出了新型的基于多目标数据生成的昂贵多目标进化算法。首先,本文提出并证明了非支配解生成定理,并基于此提出了 MDG 方法。其次,本文提出了 MPMS 框架,并结合 MDG 提出了 EMOEA-MDG 算法。此外,本文也对算法进行了复杂度分析。通过在两个著名测试集的 16 个问题上进行实验分析,并与现有的五个前沿算法进行对比,验证了本文提出的算法的性能。实验结果表明提出的算法能在 62.5% 的问题上取得比所有对比算法都更好的性能指标值,并且具有较低的时间消耗。

在未来的研究中,本文提出的昂贵多目标优化算法将会被扩展并应用到各类智能优化应用问题之中。同时,我们将会面向现实世界中不断变大的问题规模和日益复杂的问题特性,设计搜索能力更强和计算效率更高的昂贵优化算法,为解决各种关键领域应用研发和技术突破提供助力。

致 谢 感谢给本文提出意见和建议的各位专家,使得本文的质量得到极大提升!

参 考 文 献

- [1] Jin Y, Wang H, Chugh T, et al. Data-driven evolutionary optimization: An overview and case studies. *IEEE Transactions on Evolutionary Computation*, 2019, 23(3): 442-458
- [2] Zhan Z H, Shi L, Tan K C, Zhang J. A survey on evolutionary computation for complex continuous optimization. *Artificial Intelligence Review*, 2022, 55: 59-110
- [3] Li J Y, Zhan Z H, Zhang J. Evolutionary computation for expensive optimization: A survey. *Machine Intelligence Research*, 2022, 19(1): 3-23
- [4] Liu Ruo-Chen, Li Jian-Xia, Liu Jing, Jiao Li-Cheng. A survey on dynamic multi-objective optimization. *Chinese Journal of Computers*, 2020, 43(7): 1246-1278(in Chinese)
(刘若辰, 李建霞, 刘静, 焦李成. 动态多目标优化研究综述. *计算机学报*, 2020, 43(7): 1246-1278)
- [5] Wang Li-Ping, Ren Yu, Qiu Qing-Cang, Qiu Fei-Yue. Survey on performance indicators for multi-objective evolutionary algorithm. *Chinese Journal of Computers*, 2021, 44(8): 1590-1619(in Chinese)
(王丽萍, 任宇, 邱启仓, 邱飞岳. 多目标进化算法性能评价指标研究综述. *计算机学报*, 2021, 44(8): 1590-1619)
- [6] Wang Feng, Zhang Heng, Han Meng-Chen, Xing Li-Ning. Co-evolution based mixed-variable multi-objective particle swarm optimization for UAV cooperative multi-task allocation problem. *Chinese Journal of Computers*, 2021, 44(10): 1967-1983(in Chinese)
(王峰, 张衡, 韩孟臣, 邢立宁. 基于协同进化的混合变量多目标粒子群优化算法求解无人机协同多任务分配问题. *计算机学报*, 2021, 44(10): 1967-1983)
- [7] Wu S H, Zhan Z H, Zhang J. SAFE: Scale-adaptive fitness evaluation method for expensive optimization problems. *IEEE Transactions on Evolutionary Computation*, 2021, 25(3): 478-491
- [8] Li J Y, Zhan Z H, Tan K C, et al. A meta-knowledge transfer-based differential evolution for multitask optimization. *IEEE Transactions on Evolutionary Computation*, 2021, DOI: 10.1109/tevc.2021.3131236
- [9] Wang Li-Ping, Feng Mei-Ling, Qiu Qi-Cang, et al. Survey on preference-based multi-objective evolutionary algorithm. *Chinese Journal of Computers*, 2019, 42(6): 1289-1315(in Chinese)
(王丽萍, 丰美玲, 邱启仓等. 偏好多目标进化算法研究综述. *计算机学报*, 2019, 42(6): 1289-1315)
- [10] Fan Tian-Tian, Xu Lei, Chen Lin. Recommending similar bug reports based on multi-targets optimization algorithm NSGA-II. *Chinese Journal of Computers*, 2019, 42(10): 2175-2189(in Chinese)
(樊田田, 许蕾, 陈林. 基于多目标优化算法 NSGA-II 推荐相似缺陷报告. *计算机学报*, 2019, 42(10): 2175-2189)
- [11] Wang Li-Ping, Zhang Ming-Lei, Qiu Fei-Yue, Jiang Bo. Many-objective optimization algorithm with preference based on the angle penalty distance elite selection Strategy. *Chinese Journal of Computers*, 2018, 41(1): 236-253(in Chinese)
(王丽萍, 章鸣雷, 邱飞岳, 江波. 基于角度惩罚距离精英选择策略的偏好高维多目标优化算法. *计算机学报*, 2018, 41(1): 236-253)
- [12] Xie Cheng-Wang, Yu Wei-Wei, Guo Hua, et al. DAV-MOEA: A many-objective evolutionary algorithm adopting dynamic angle vector based dominance relation. *Chinese Journal of Computers*, 2022, 45(2): 317-333(in Chinese)
(谢承旺, 余伟伟, 郭华等. DAV-MOEA: 一种采用动态角度向量支配关系的高维多目标进化算法. *计算机学报*, 2022, 45(2): 317-333)
- [13] Koziel S, Leifsson L. Multi-level CFD-based airfoil shape optimization with automated low-fidelity model selection. *Procedia Computer Science*, 2013, 18(1): 889-898
- [14] Gu Qing-Hua, Zhou Yu-Feng, Li Xue-Xian, Ruan Shun-Ling. Expensive many-objective evolutionary algorithm based on radial space division. *Acta Automatica Sinica*, 2021, 48(10): 2564-2584(in Chinese)
(顾清华, 周煜丰, 李学现, 阮顺领. 基于径向空间划分的昂贵多目标进化算法. *自动化学报*, 2021, 48(10): 2564-2584)
- [15] Sun Zhe-Ren, Huang Yu-Hua, Chen Zhi-Yuan. Diversity based surrogate-assisted evolutionary algorithm for expensive multiobjective optimization problem. *Journal of Software*, 2021, 32(12): 3814-3828(in Chinese)

- (孙哲人, 黄玉划, 陈志远. 面向多目标优化的多样性代理辅助进化算法. 软件学报, 2021, 32(12): 3814-3828)
- [16] Ponweiser W, Wagner T, Biermann D, et al. Multiobjective optimization on a limited budget of evaluations using model-assisted S-metric selection//Proceedings of the International Conference on Parallel Problem Solving from Nature. Berlin, Germany, 2008: 784-794
- [17] Zhang Q, Liu W, Tsang E, Virginas B. Expensive multiobjective optimization by MOEA/D with Gaussian process model. IEEE Transactions on Evolutionary Computation, 2010, 14(3): 456-474
- [18] Li J Y, Zhan Z H, Wang H, Zhang J. Data-driven evolutionary algorithm with perturbation-based ensemble surrogates. IEEE Transactions on Cybernetics, 2021, 51(8): 3925-3937
- [19] Li J Y, Zhan Z H, Wang C, et al. Boosting data-driven evolutionary algorithm with localized data generation. IEEE Transactions on Evolutionary Computation, 2020, 24(5): 923-937
- [20] Liu X F, Zhan Z H, Zhang J. Resource-aware distributed differential evolution for training expensive neural-network-based controller in power electronic circuit. IEEE Transactions on Neural Networks and Learning Systems, 2021, early access, DOI:10.1109/TNNLS.2021.3075205
- [21] Li J Y, Zhan Z H, Xu J, et al. Surrogate-assisted hybrid-model estimation of distribution algorithm for mixed-variable hyperparameters optimization in convolutional neural networks. IEEE Transactions on Neural Networks and Learning System, 2021, DOI: 10.1109/TNNLS.2021.3106399
- [22] Pan L, He C, Tian Y, et al. A classification-based surrogate-assisted evolutionary algorithm for expensive many-objective optimization. IEEE Transactions on Evolutionary Computation, 2019, 23(1): 74-88
- [23] Guo D, Wang X, Gao K, et al. Evolutionary optimization of high-dimensional multiobjective and many-objective expensive problems assisted by a dropout neural network. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, DOI: 10.1109/TSMC.2020.3044418
- [24] Wang H, Jin Y. A random forest assisted evolutionary algorithm for data-driven constrained multi-objective combinatorial optimization of trauma systems. IEEE Transactions on Cybernetics, 2020, 50(2): 536-549
- [25] Guo D, Jin Y, Ding J, Chai T. Heterogeneous ensemble-based infill criterion for evolutionary multiobjective optimization of expensive problems. IEEE Transactions on Cybernetics, 2019, 49(3): 1012-1025
- [26] Cheng R, Jin Y, Olhofer M, Sendhoff B. A reference vector guided evolutionary algorithm for many-objective optimization. IEEE Transactions on Evolutionary Computation, 2016, 20(5): 773-791
- [27] Chugh T, Jin Y, Miettinen K, et al. A surrogate-assisted reference vector guided evolutionary algorithm for computationally expensive many-objective optimization. IEEE Transactions on Evolutionary Computation, 2018, 22(1): 129-142
- [28] Wang X, Jin Y, Schmitt S, Olhofer M. An adaptive Bayesian approach to surrogate-assisted evolutionary multi-objective optimization. Information Science, 2020, 519: 317-331
- [29] Zhan Z H, Li J, Cao J, et al. Multiple populations for multiple objectives: A coevolutionary technique for solving multiobjective optimization problems. IEEE Transactions on Cybernetics, 2013, 43(2): 445-463
- [30] Liu S C, Chen Z G, Zhan Z H, et al. Many-objective job-shop scheduling: A multiple populations for multiple objectives-based genetic algorithm approach. IEEE Transactions on Cybernetics, 2021, DOI: 10.1109/TCYB.2021.3102642
- [31] Chen Z G, Zhan Z H, Lin Y, et al. Multiobjective cloud workflow scheduling: A multiple populations ant colony system approach. IEEE Transactions on Cybernetics, 2019, 49(8): 2912-2926
- [32] Zhang X, Zhan Z H, Fang W, et al. Multi population ant colony system with knowledge based local searches for multiobjective supply chain configuration. IEEE Transactions on Evolutionary Computation, 2022, 26(3): 512-526
- [33] Li J Y, Deng X Y, Zhan Z H, et al. A multipopulation multiobjective ant colony system considering travel and prevention costs for vehicle routing in COVID-19-like epidemics. IEEE Transactions on Intelligent Transportation Systems, 2022, DOI: 10.1109/TITS.2022.3180760
- [34] Zhou S Z, Zhan Z H, Chen Z G, et al. A multi-objective ant colony system algorithm for airline crew rostering problem with fairness and satisfaction. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(11): 6784-6798
- [35] Knowles J. ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. IEEE Transactions on Evolutionary Computation, 2006, 10(1): 50-66
- [36] Cai Xin-Ye, Ma Zhong-Yu, Zhang Feng, et al. Adaptive multitask with multipopulation-based cooperative search for expensive multiobjective optimization problems. Chinese Journal of Computers, 2021, 44(9): 1934-1948(in Chinese) (蔡昕烨, 马中雨, 张峰等. 基于自适应分解的多任务协作型昂贵多目标优化算法. 计算机学报, 2021, 44(9): 1934-1948)
- [37] Luo J, Gupta A, Ong Y S, Wang Z. Evolutionary optimization of expensive multiobjective problems with co-sub-Pareto front Gaussian process surrogates. IEEE Transactions on Cybernetics, 2019, 49(5): 1708-1721
- [38] Yuan Y, Banzhaf W. Expensive multi-objective evolutionary optimization assisted by dominance prediction. IEEE Transactions on Evolutionary Computation, 2022, 26(1): 159-173
- [39] Tian Y, Cheng R, Zhang X, et al. An indicator based multi-objective evolutionary algorithm with reference point adaptation for better versatility. IEEE Transactions on Evolutionary Computation, 2018, 22(5): 609-622
- [40] Song Z, Wang H, He C, Jin Y. A Kriging-assisted two-arch evolutionary algorithm for expensive many-objective optimization. IEEE Transactions on Evolutionary Computation, 2021, 25(6): 1013-1027

- [41] Liu Jian-Chang, Zhao Yang-Jie, Li Fei, et al. Expensive multi-objective evolutionary algorithm based on R2 index. *Control and Decision*, 2020, 35(4): 823-832(in Chinese) (刘建昌, 赵阳杰, 李飞等. 基于 R2 指标的昂贵多目标进化算法. *控制与决策*, 2020, 35(4): 823-832)
- [42] Habib A, et al. A multiple surrogate assisted decomposition-based evolutionary algorithm for expensive multi/many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 2019, 23(6): 1000-1014
- [43] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 2002, 6(2): 182-197
- [44] Deb K, Thiele L, Laumanns M, Zitzler E. Scalable test problems for evolutionary multiobjective optimization//Abraham A, Jain L, Goldberg R, eds. *Evolutionary Multi-Criterion Optimization*. London: Springer, 2005: 105-145
- [45] Huband S, Hingston P, Barone L, While L. A review of multiobjective test problems and a scalable test problem toolkit. *IEEE Transactions on Evolutionary Computation*, 2006, 10(5): 477-506
- [46] Tian Y, Cheng R, Zhang X, Jin Y. PlatEMO: A MATLAB platform for evolutionary multi-objective optimization [educational forum]. *IEEE Computational Intelligence Magazine*, 2017, 12(4): 73-67



LI Jian-Yu, Ph. D. candidate. His main research interests include artificial intelligence, evolutionary computation, and expensive optimization.

ZHAN Zhi-Hui, Ph. D., professor. His main research interests include artificial intelligence, evolutionary computation, and swarm intelligence.

Background

Expensive multi-objective optimization problem are those complex optimization problems with multiple objectives that are not only conflicted with each other and but also computationally expensive, which exist in various fields, such as and. Therefore, the research into how to solve such kind of problems has become an increasingly hot and significant research topic worldwide. The key issue in developing an efficient expensive multi-objective optimization algorithm lies in how to efficiently and simultaneously handle the diversity and convergence challenges, i. e., how to efficiently obtain non-dominated solutions with good diversity and high quality under limited computational resource.

To address the diversity and convergence challenges, this paper proposes a novel multi-objective data generation-based expensive multi-objective evolutionary algorithm. There are three contributions and innovations of this paper when compared with existing work. First, this paper puts forward and proves the non-dominated solution generation theorem, and then proposes a multi-objective data generation method based on the theorem, so as to obtain more non-dominated solutions more efficiently for improving the solution diversity. Second, this paper proposes a multiple population multiple surrogates framework that co-evolves multiple populations to

efficiently optimize multiple surrogates that are built for multiple real expensive objectives respectively. Third, based on the above proposed method and framework, this paper proposes an expensive multi-objective evolutionary algorithm based on multi-objective data generation for efficiently solving the expensive multi-objective optimization problem. Extensive experimental results have shown the great superiority of the proposed algorithm and methods over compared state-of-the-art algorithms.

Our research team has been devoted to the research of evolutionary computation algorithms and their applications in real-world optimization problems. We have published more than 160 papers in international journals and conferences, including more than 50 IEEE Transactions papers and more than 10 ESI Highly Cited Papers. Specifically, we have published more than 10 IEEE Transaction papers and 1 ESI Highly Cited Paper in the related field of expensive optimization and multi-objective optimization.

This work was supported in part by the National Key Research and Development Program of China under Grant 2019YFB2102102 and in part by the National Natural Science Foundations of China under Grant 62176094.