

Trend Prediction Model Based Multi-Objective Estimation of Distribution Algorithm

Zhongqiang Huang, Min Jiang

Fujian Key Lab of the Brain-Like Intelligent Systems, Xiamen University (XMU), Xiamen Fujian
Email: usepr@stu.xmu.edu.cn, minjiang@xmu.edu.cn

Received: Mar. 11th, 2016; accepted: Mar. 26th, 2016; published: Mar. 30th, 2016

Copyright © 2016 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Multi-objective optimization problems exist widely in real world applications. Traditional evolutionary algorithms usually employ individual-based evolution strategies to solve these optimization problems, leading to low convergence rate, strong dependency on population size and poor results. As a meta-heuristic algorithm, the Estimation of Distribution Algorithm (EDA) combines the statistical machine learning with population evolution model and has attracted a wide spread attention. In this paper, we proposed a trend-prediction-model (TPM) based EDA method, called TPM-EDA, to solve multi-objective problems. The characteristic of TPM is that it effectively utilizes the historic information generated in evolutionary process to predict the trend of particles, so as to promote the search speed for finding Pareto-optimal front and the search ability of algorithm. Meanwhile, the sparseness is applied in our algorithm to control the sampling frequencies of individuals for the purpose of achieving the diversity of population. We compared our method with multiple existing EDA algorithms on 6 different test instances. The experimental results proved the effectiveness of our method.

Keywords

Multi-Objective Optimization, Estimation of Distribution Algorithm, Trend Prediction Model

基于趋势预测模型的多目标分布估计算法

黄忠强, 江敏

厦门大学福建省仿脑智能系统重点实验室, 福建 厦门
Email: usepr@stu.xmu.edu.cn, minjiang@xmu.edu.cn

收稿日期: 2016年3月11日; 录用日期: 2016年3月26日; 发布日期: 2016年3月30日

摘要

多目标优化问题广泛存在于现实世界的应用当中。传统的基于个体进化策略的进化算法在处理这些优化问题时往往收敛速度慢、严格依赖于种群大小而且效果不大理想。分布估计算法作为元启发式(meta-heuristics)方法的一种, 将统计机器学习同群体进化模式相结合, 引起了学者的广泛关注。在这篇文章中, 我们提出了一种基于趋势预测模型(TPM)的分布估计算法, TPM-EDA, 用于解决多目标优化问题。其特点在于有效地利用了群体进化过程的历史信息来预测粒子运动的趋势, 从而加速了查找最优Pareto前沿面的过程, 提升了算法的搜索能力。与此同时, 通过引入稀疏度来控制个体的采样频率, 来实现种群的多样性。我们在6个不同的测试函数上, 对TPM-EDA和多种已有的EDA算法进行了对比性试验。实验结果表明了TPM-EDA方法的有效性。

关键词

多目标优化, 分布估计算法, 趋势预测模型

1. 引言

现实世界中的很多应用都与多目标优化问题存在着紧密的联系, 例如购车时希望能买到性能好、价格低廉同时又省油的车。这种紧密的联系也意味着高效的多目标问题的求解方法, 能够带来重要的经济利益。因此, 对多目标优化问题的研究, 尤其是设计通用而且高效的算法受到了相关领域的广泛关注[1]-[3]。

大量的多目标优化问题难以通过解析方式求得其最优解, 这也就使得设计不同的启发式算法成为解决多目标优化问题的主要方法之一[4] [5]。分布估计算法[6]作为元启发式(meta-heuristics)方法的一种, 其主要特点是将统计机器学习的方法同群体进化模式进行结合。与传统的遗传算法不同, 在分布估计算法中, 没有交叉、变异等操作, 取而代之的是概率模型的学习和采样。EDA 算法往往通过统计学习的手段从宏观的角度为解的分布建立一个概率模型, 通过不同的方法在该模型上采样来得到下一代种群, 如此反复进行, 实现种群的进化。由于在优化的过程中显式的使用概率模型, EDA 算法能够提供其他元启发方法所不能提供的优点。

近些年来, 利用 EDA 算法来解决多目标优化问题, 已经取得了很好的结果。Khan 等[7]提出的多目标贝叶斯优化算法(mBOA)将贝叶斯优化算法同非支配排序算法相结合, 利用 NSGA-II 中的选择方法来沿着 Pareto 前沿面寻找解。这种方法的特点在于通过提供选择压力来确保 Exploitation 和 Exploration 之间取得某种平衡; Pelikan 等[8]提出的多目标分层 BOA (mohBOA), 也借鉴了 NSGA-II 和 mBOA 中的类似过程来选择可能的候选解, 但是不同之处在于在选出了候选解之后, mohBOA 结合了 k-means 聚类算法来获得有希望的解的聚类, 并使用限制的二进制锦标赛法来将新获得的聚类来与旧的种群集成, 来产生下一代种群。实验结果也都证明了这种方法的有效性。Shah 等人证明[9], ϵ -hBOA, 一种基于 ϵ 占有的分层贝叶斯优化算法, 在求解多目标 d 维背包问题是较其他两种进化算法有一定优势。Karshenas 等人提出的 MBN-EDA [10]建立了关于决策变量和目标变量间关系的多维贝叶斯网络。MBN-EDA 不仅捕获了变量间的依赖关系, 同时也刻画了目标与目标以及目标与变量间的相关关系。然而训练一个贝叶斯网络

是复杂而耗时的,会造成巨大开销。除了基于贝叶斯网络的 EDA 方法,还有基于规则化模型的 RM-MEDA [11]、基于 GNG 网络的 MB-GNG [12]、基于 Parzen 窗口的 MOPEDA [13]、基于方向模型的 IM-MOEA [14] 等。

然而,在相当多的情况下,尤其是对于多目标优化问题而言,传统 EDA 算法所用来建模的概率模型很难描述真实前沿面的概率分布。当模型偏差较大时,将直接导致 EDA 算法的失效。一种有效的解决方式是直接从某个带启发式信息的分布上进行采样然后过滤个体,来替代传统 EDA 的建模和采样过程。如 Luhan Zhou 等人在[15]中所使用的策略,它强调了生成个体和滤除个体之间的配合,展现出某些传统 EDA 算法所不具备的优势。我们的方法正是基于这种策略来不断地产生粒子然后过滤粒子,来实现种群的进化。

在这篇论文中,我们提出了一种称为 TPM-EDA 的算法用于解决多目标优化问题,我们工作的主要贡献在如下两点:

- 利用稀疏度来控制个体的采样频率,使得种群能均匀地分布在 Pareto 前沿面上,实现了种群的多样性;
- TPM-EDA 提出了一种不同的采样方法来生成下一代种群。其在已经选择的个体的基础上,构造一个基于历史信息的发展趋势预测模型,并在采样过程中使用该模型预测下一代粒子的位置,从而加速寻找前沿面的速度。

本文的文章结构如下。在第二节,介绍了多目标优化问题的相关概念以及分布估计算法。第三节先介绍了 TPM-EDA 的算法框架,而后详细描述了我们所提出的趋势预测模型以及 TPM-EDA 算法。第四节是实验部分,包括测试样例、参数设置以及实验结果。最后是总结和展望。

2. 多目标优化方法与进化计算

2.1. 多目标优化问题定义

在这篇论文中,我们考虑无约束的多目标优化问题。一个无约束的多目标问题定义为:

$$\text{minimize } F(x) = (f_1(x), \dots, f_m(x))$$

其中, $x = (x_1, \dots, x_n)$ 是 n 维决策变量, f_i 是第 i 个目标函数。

定义 1. (Pareto 支配)我们称一个解 x_1 支配另一个解 x_2 , 记作 $x_1 \succ x_2$, 如果

$$\begin{cases} f_i(x_1) \leq f_i(x_2), \forall i = 1, \dots, m \\ f_i(x_1) < f_i(x_2), \forall i = 1, \dots, m \end{cases}$$

定义 2. (Pareto 最优解集, POS)我们将一个多目标问题的解集记为 POS,

$$\text{POS} = \{x^* \mid \nexists x, x \succ x^*\}$$

定义 3. (Pareto 前沿面, POF)我们将一个多目标问题的前沿面记为 POF,

$$\text{POF} = \{F(x) \mid x^* \in \text{POS}\}$$

与单目标优化问题寻找一个最优解不同的是,多目标优化问题需要寻找一个最优的解集,即上述定义的 POS,使得最优解集中的解不会被其他解支配。多目标优化算法的目的便是寻找到一个最接近理论前沿面的 Pareto 前沿面及其对应的解集。

2.2. 分布估计算法 EDA

EDA 算法(Estimation of Distribution Algorithms) [6] [16]-[18]将统计机器学习与进化计算相结合,通过

将概率模型引入进化框架中来对问题进行优化。与基于个体的进化策略不同的是, EDA 是基于种群优化的, 它通过反复地对候选解所处的空间进行建模和采样来搜索空间。由于 EDA 的优化过程是概率分布的反复更新, 使得它相比较传统进化算法有如下优势: 1) 自适应算子, 它们可以根据要解决的问题来调整算子; 2) 清晰的问题结构, 它们提供了问题解的具体概率模型, 使得模型的分析 and 信息挖掘变得容易; 3) 先验概率的存在, EDA 可以加入先验概率到模型中; 4) 降低了内存需要, EDA 只需记录概率模型参数而无需直接保存种群, 降低了对内存的需求。EDAs 已经成功应用在多目标背包问题[19], 经济调度[20], 蛋白质结构预测[21]等上。

近些年来, 分布估计算法发展了很多不同的具体实现方法, 其差别往往在于采用了不同的概率模型生成方法和不同的采样方法。尽管不同方法的设计思想有很大的不同, 但是这些方法都可以归纳为下面两个主要步骤:

- 首先通过某种方法来构建用于描述解空间的概率模型。EDA 算法往往会通过对种群的评估, 选择优秀的个体集合, 然后采用统计学习等手段构造一个描述当前解集的概率模型。我们这这位概率模型的生成。
- 在概率模型的基础上, 利用不同的采样方式来产生新的种群。我们称此步骤为采样方法。

EDA 算法的大体流程如图 1 所示。

大多数的 EDA 算法都是通过显式地构建一个概率模型, 并在这个模型上进行采样来得到新的种群。然而要准确地构建概率模型是一件很困难的事情, 错误的概率模型将直接导致新得到的种群质量不高, 因此很难找到问题的解。一种替代的方法是隐式的建立概率模型, 通过不断地产生和过滤个体达到。我们的思路是首先根据稀疏度选择待扩展的个体, 然后进行采样, 并通过筛选非支配解集来得到优势种群。另外一方面, 我们认同 Claudio Rossi 等人的观点[22], 即充分利用某些信息, 将能够使得算法提高对变化的前沿面的追踪能力。作者使用了卡尔曼滤波来预测变化的最优点, 而我们则认为不仅简单的预测机制对于多目标优化问题而言至关重要, 于此同时概率知识, 而不是精确值, 能够给我们更多的帮助。因此, 我们提出了利用基于运动趋势预测的机制来对前沿面进行预测。

3. TPM-EDA 方法

在这一节的开始部分, 我们首先定义相关的术语, 然后给出 TPM-EDA 方法的算法框架。接下来我们将更加详细地介绍 TPM-EDA 的两个组成部分: 待扩展个体的选择过程和趋势预测模型。前者用于选取待扩展的个体, 而后者则是利用历史信息来预测下一代种群。在介绍完这两者之后, 我们将给出 TPM-EDA 具体的算法描述。

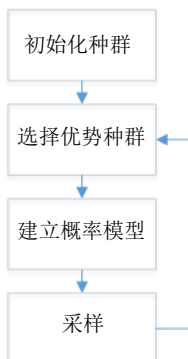


Figure 1. Flowchart of estimation of distribution algorithm
图 1. 分布估计算法流程图

3.1. TPM-EDA 算法框架

定义 4. (粒子) 粒子是一个二元组 $p = \langle x, d \rangle$, 其中 x 表示的是多目标优化问题的解, 也叫做个体; 而 d 是 n 维空间中的单位向量, 表示粒子的方向。

在 TPM-EDA 算法中, 每次采样时都挑选一个待扩展的粒子, 然后在这个粒子所代表的一个特定分布上进行采样。下面, 我们结合图 2(a)~(e) 来简要介绍一下 TPM-EDA 的算法框架。在某一个时刻, TPM-EDA 首先产生若干个粒子(图 2(a)), 并计算对应的 POS (图 2(b))。随后计算每个粒子的稀疏度, 并根据稀疏度来选择粒子, 其原则是稀疏度较高的粒子将会有更高的概率被选中(图 2(c))。将趋势预测模型作用到所选择的粒子上, 得到一系列新的解(图 2(d)), 并由此得到新的 POS 和 POF (图 2(e))。以此不断循环, 直到停止条件满足。

3.2. 选择过程

TPM-EDA 的做法是在采样阶段挑选那些更“重要”的粒子, 从而保持粒子的多样性和前沿面的分布性。我们的做法是在那些稀疏的个体附近产生更多的子代, 而密集的区域只产生少量的个体, 即以稀疏度为权重来选择个体。假设当前群体为 $P = \{p_1, \dots, p_n\}$, 每个个体的拥挤度为其他粒子对其的相互作用之和。距离更近的个体相互作用应该更高, 因此我们定义粒子 p_k 的拥挤度为:

$$\text{Crowdedness}(p_k) = \sum_{i=1}^n \exp\left(-\frac{\|x_k - x_i\|}{h}\right)$$

其中, 参数 h 表示带宽。个体 p_k 的稀疏度则定义为拥挤度的倒数:

$$\text{Sparseness}(p_k) = \frac{1}{\text{Crowdedness}(p_k)}$$

每次采样时我们依据粒子的 Sparseness 值进行轮盘赌法。可以看出当粒子处在较稀疏的位置时, 被采样的概率较高, 能在其领域内产生更多的子代; 而较密集的地方产生子代的概率小。这样, 我们调节了所获得前沿面的稀疏程度, 从而能够保持粒子的多样性, 使前沿面更具有分布性。

3.3. 趋势预测模型

在上述选择过程挑选完待扩展的粒子后, 我们需要在这粒子所代表的某个特定分布上进行采样。在物理上, 如果知道了一个物体的运动方向 u , 那么下一时刻的位移方向 s 便等于 u 。可以写成:

$$\text{Prop}(s) = \begin{cases} 1, & \text{if } s = u \\ 0, & \text{otherwise} \end{cases}$$

与这种简单的线性预测不同的是, 我们希望给物体的运动添加一些概率性, 使得速度只是提供给物体一个可能的运动趋势而非直接确定下一时刻运动方向。这样的运动趋势可以很好地引导粒子进行更新, 从而更快、更精确地逼近前沿面。

定义 5. (趋势预测模型, TPM)

任意给定向量 $u \in \mathbb{R}^n$ 和距离度量函数 $d(\cdot, \cdot)$ 。我们把 $D \subset \mathbb{R}^n$ 上的一个概率分布 P 叫做趋势预测模型, 如果其满足:

$$\forall t_1, t_2, d(u, t_1) \geq d(u, t_2) \Rightarrow P(t_1) \leq P(t_2)$$

这样的分布刻画了粒子的运动趋势: 将以较大的概率沿着原速度 u 方向, 而又允许粒子以较小的概率违背这个方向。

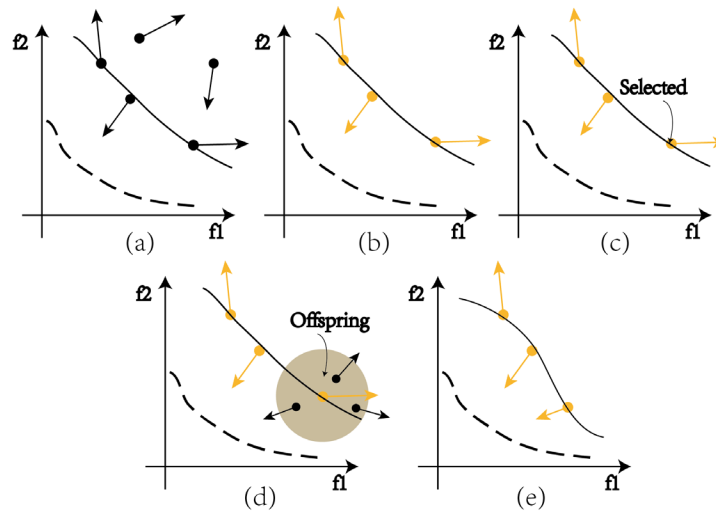


Figure 2. Algorithm framework of TPM-EDA
图 2. TPM-EDA 算法框架

图 3 展示了这样一个 TPM 实例。其中黑点是依据稀疏度被选择后的粒子， u 是其运动方向。如果 u 所代表的 TPM 模型以 \cos 距离为距离度量方式，定义域为单位球面，则因为 v_1 相比较 v_2 更靠近 u （反应在 \cos 距离上是二者夹角更小），所以 v_1 被选择到的概率更高。类似的， v_3 的概率比 v_2 这概率更低，而 v_4 最不可能被采样到。

在我们的算法当中，每个粒子运动的启发式信息，产生它的父节点指向该节点的方向向量，被作为运动方向记录下来。我们希望设计一个趋势预测模型使得采样出的向量离这个运动方向 u 越近则概率越高。为此，我们设计了一个特定的 TPM 模型 TPM_h 。设定该模型距离度量方式为 \cos 距离，定义域 D 为单位球面，位移方向与原速度的夹角从均值为 0，方差为带宽 $1/h$ 的高斯分布上进行采样，而粒子的步长则由均值为 0，方差为 h 的高斯分布上产生。由于 D 是单位球面，所以 TPM_h 其实是关于方向向量的分布， \cos 距离则刻画了采样出来的位移方向与原速度之间的夹角大小。算法 `sample` 描述了这个采样过程的具体实现。

`sample(PS, h, N)`

1. $samples \leftarrow \emptyset$;
 2. For $i = 1$ to N
 3. 计算粒子 PS_i 的稀疏度;
 4. End For
 5. For $i = 1$ to N
 6. 依据稀疏度，使用轮盘赌法随机选一个例子 p ;
 7. 从分布 $\text{Norm}(0, 1/h)$ 中采样出一个角度 θ ;
 8. 随机选择一个与粒子速度 v 夹角为 θ 的方向向量 u ;
 9. 从分布 $\text{Norm}(0, h)$ 中采样出一个步长 λ ;
 10. 添加新的粒子 $p + \lambda u$ 进 $samples$;
 11. End For
 12. Return $samples$;
-

因为每个粒子都带有能够刻画运动趋势的方向向量, 因此能够准确预测前沿面的推进方向。为了实现时间维度上的开发与扩展平衡, 更具体地, 要实现在算法前期具有较高的搜索速度, 在后期具有较细致的搜索能力, 我们引入趋势度的概念。通过控制算法不同阶段趋势度来调节这个平衡。注意到, 当 h 趋于 0 时, TPM_h 趋于均匀分布。均匀分布可以看作没有任何趋势的分布, 因此, 我们定义 TPM_h 的趋势度为其与均匀分布的距离。

定义 6. (趋势度, TD) 定义趋势预测模型 TPM_h 的趋势度为:

$$\text{TD}(\text{TPM}_h) = \text{dis}(\text{TPM}_h, \text{Uniform})$$

其中 dis 为分布间距离的度量方式, Uniform 表示均匀分布。令算法中的参数带宽 h 随着迭代不断降低。基于上述定义, TPM_h 模型的趋势度将随着时间递减。算法运行阶段前期趋势度大, 粒子会最大可能沿着速度方向前进, 从而较快逼近前沿面; 后期趋势度不断降低, 来达到局部的、更细致地搜索。

我们需要设定个停止条件来让算法终止。定义如下指标来衡量两个前沿面 P_1 、 P_2 之间的差距:

$$\Delta(P_1, P_2) = \frac{\sum_{p \in P_1} \min_{q \in P_2} \|p - q\| + \sum_{p \in P_2} \min_{q \in P_1} \|p - q\|}{|P_1| + |P_2|}$$

该条件描述了离前沿面的临近程度, 若 P_1 、 P_2 一致, 则 $\Delta(P_1, P_2) = 0$ 。因为参数带宽 h 随着迭代不断降低, 粒子运动的步长将不断趋于 0, 因此以该条件作为 TPM-EDA 算法的停止条件时, 算法是可以停止的。

3.4. TPM-EDA 算法

综合上述的选择过程以及趋势预测模型, 我们给出 TPM-EDA 算法的具体实现。

算法 TPM-EDA

输入: C : 种群限制, N : 子代数目

输出: POS: 最优解集

1. 随机初始化 N 个个体;
 2. 初始化 POS、POF;
 3. 初始化衰减系数 r 和带宽 h ;
 4. While $\Delta(\text{POF}, \text{POF}') > \epsilon$
 5. $\text{samples} = \text{sample}(\text{POS}, 1/h, N)$;
 6. 利用 samples 更新 POS、POF;
 7. 反复删除最密集的个体直到 $|\text{POS}| \leq C$;
 8. $h = h * r$;
 9. End While
 10. 打印 POS;
-

4. 实验

我们将 GM-EDA 和其他的 EDA 方法进行了对比性实验, 画出了它们解出的前沿面, 并利用 IGD 指标来对各种算法的求得的解的收敛性和多样性进行比较。The inverted generational distance (IGD) [25] 被用来衡量多目标算法的性能。令 P^* 为均匀分布在最优前沿面上的点集, 而 P 是算法求得的前沿面点集。定

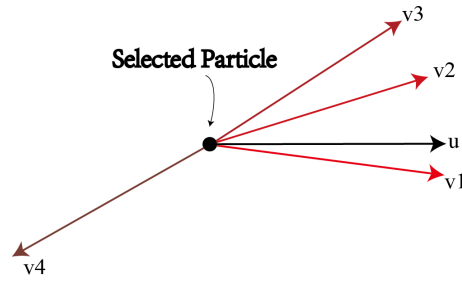


Figure 3. An example of TPM distribution
图 3. 一个 TPM 实例

义 IGD 指标为:

$$\text{IGD}(P^*, P) = \frac{\sum_{v^* \in P^*} \min_{v \in P} \|v^* - v\|}{\|P^*\|}$$

要使得 IGD 值尽量小, 需要 P 尽量逼近 P^* 。因此, 该定义刻画了算法求得的前沿面与最优前沿面之间的相近程度。

4.1. 测试样例和参数设置

在这篇文章中, 我们将所提出的方法与其他四种 EDA 算法相比较, 包括 MBN-EDA [10], RM-MEDA [11]和 EGNA [23]。MBN-EDA 方法基于联合概率模型, 它通过建立多维的贝叶斯网络来捕获决策变量和目标变量间的关系。RM-MEDA 是基于规则模型的多目标优化算法。它假定 Pareto 解集处在分段连续的 $m-1$ 维流形空间中并使用局部的主成分分析来为候选区域进行建模。EGNA 是使用高斯网络来为决策变量建立概率模型的一种简单 EDA 方法, 但它并不考虑目标的信息。

在实验中, 我们使用了 6 个测试样例[24], 包括 FDA4、FDA5、DIMP2、dMOP2、dMOP3 和 HE2。这些测试函数具有不同的特性, 它们的理论前沿面具有不同的形状, 比如连续型的、分离型的等。FDA4 和 FDA5 是 3 目标函数而其他是两目标函数。

实验中的 EDA 算法都遵循相同的环境设置: 每次迭代至多含有 100 个个体, 最多迭代次数 200。各个算法的参数与它们各自的文章中一致: 对于 MBN-EDA, 使用利润增益排序(Profit gain ordering)作为排序方法[10]; RM-MEDA 中, 局部 PCA 聚类的数目被设置为 5 [11]; EGNA 中截断选择被用作选择方法 [23]。

4.2. 结果

由于篇幅有限, 我们只选取部分算法和测试函数来画出他们所获得的前沿面的情况。图 4 中画出了三个 EDA 方法求解得到的前沿面情况。红色的点表示算法求得的个体, 3D 线框展示的是理论的前沿面。从图中可以看出, TPM-EDA 方法取得了接近真实前沿面的结果。

图 5 展示的是 DIMP2 函数的情况。红色的点是算法求得的解, 蓝色实线表示理论的前沿面。所测试的三个方法在这个测试函数上表现都不大理想。但相对而言, TPM-EDA 的效果较好, 它获得的种群具有更好的分布性。

图 6 展示的是 HE2 函数的情况, 该函数的前沿面是分离的。从图中可以看出 TPM-EDA 具有最好的收敛效果, 其前沿面最贴近真实的前沿面。然而 TPM-EDA 算法缺少位于最右侧的前沿面片段中的个体, 导致它前沿面的离散程度弱于 RM-MEDA。

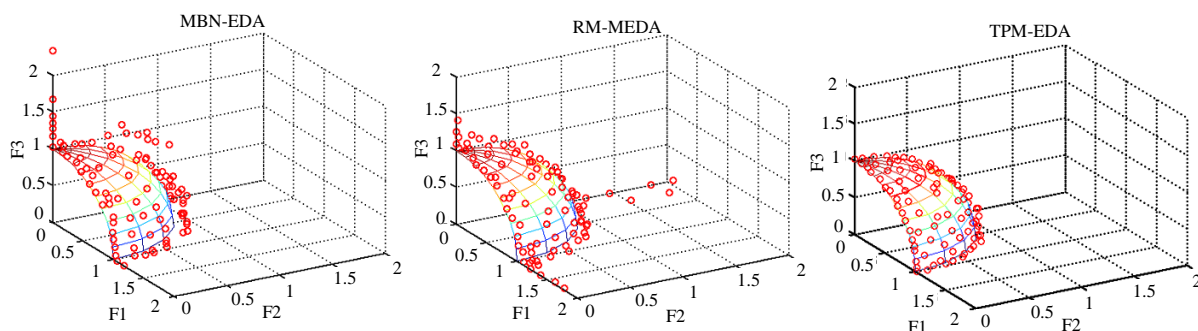


Figure 4. POF of FDA4 obtained by MBN-EDA, RM-MEDA and TPM-EDA

图 4. MBN-EDA, RM-MEDA 和 TPM-EDA 获得的 FDA4 前沿面

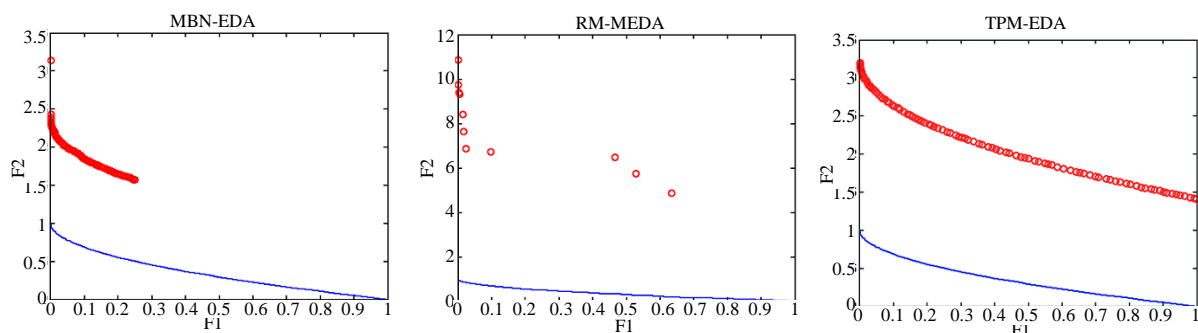


Figure 5. POF of DIMP2 obtained by MBN-EDA, RM-MEDA and TPM-EDA

图 5. MBN-EDA, RM-MEDA 和 TPM-EDA 获得的 DIMP2 前沿面

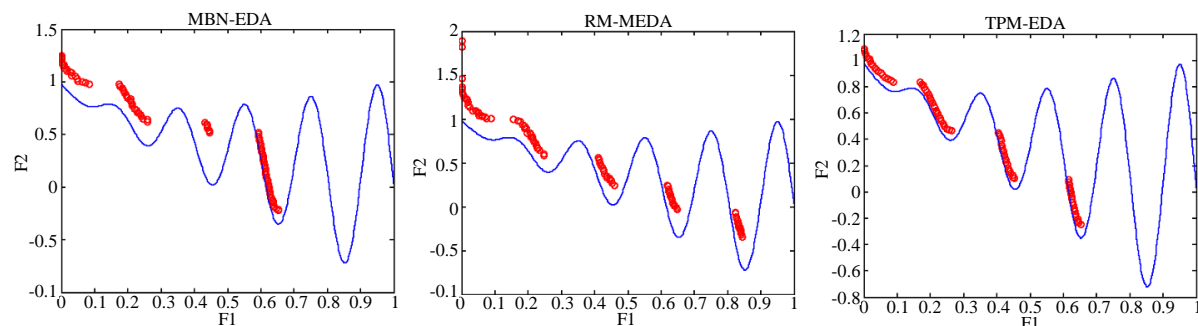


Figure 6. POF of HE2 obtained by MBN-EDA, RM-MEDA and TPM-EDA

图 6. MBN-EDA, RM-MEDA 和 TPM-EDA 获得的 HE2 前沿面

为了考察算法解的收敛性, 对每个标准函数, 我们都进行了 20 次测试。对每次测试得到的 IGD 值取平均得到算法在这个测试函数上的 IGD 度量, 所获得的结果在表 1 当中。可以看出 TPM-EDA 在大部分测试函数上都超过了其他算法。在 dMOP2 测试函数上, TPM-EDA 也处在了第二的位置。因此, 算法求得的前沿面具有较好的收敛性, 能够很好地逼近理论的最优解。

4.3. 讨论

从静态实验结果以及 IGD 指标上都可以看出 TPM-EDA 算法能取得比其他 EDA 算法更优秀的结果。基于稀疏度的选择过程以及趋势预测模型的使用使得 TPM-EDA 能够更快、更准确地逼近理论前沿面。趋势预测模型有效地利用了粒子更新过程中的历史信息, 以此预测粒子运动的趋势。

我们可以将所提出的 TPM-EDA 方法应用到很多实际应用中。机器学习中的很多问题都可以转化为

Table 1. IGD values of each method
表 1. 各个方法的 IGD 指标

IGD	MBN-EDA	RM-MEDA	EGNA	TPM-EDA
FDA4	0.43	0.44	1.43	0.32
FDA5	0.51	0.47	1.53	0.40
DIMP2	6.97	5.12	6.56	3.56
dMOP2	1.40	0.29	1.35	0.38
dMOP3	1.38	1.05	1.86	0.98
HE2	0.83	0.74	1.08	0.50

多目标优化问题, 如在深度学习中, 我们可以利用 TPM-EDA 同时优化深度神经网络的训练误差和网络结构[26]; 在迁移学习的领域适应方法中, 可以利用多目标优化算法来同时优化映射后领域的相似性以及样本的可分性[27]; 在云计算的多任务调用中找到效率高而又节能的调度方案[28]等。

5. 总结和展望

本文提出了一个新的基于趋势预测模型的分布估计算法, TPM-EDA, 用来解决多目标优化问题。TPM-EDA 通过反复的产生子代个体和筛选优势种群来促进种群的进化, 并使用方向向量这个启发式信息来预测子代可能所处的区域, 从而提升搜索的准确性。此外, 基于稀疏度的选择方法被用来调节个体的稀疏程度、控制种群的多样性, 使得获得的前沿面分布性良好。实验结果也充分证明了我们方法的有效性。

一个高效的多目标优化算法是很多问题的基础。我们后期的研究方向包括基于多目标优化的深度学习及迁移学习[26] [27], 机器人自主学习[29]-[31]以及如何利用多目标优化辅助亚符号模型的知识抽取上[32]-[36]。

基金项目

此研究工作受到国家自然科学基金项目(No. 61003014, No. 61105026 和 No. 61273338)的资助。

参考文献 (References)

- [1] Ahmadi, M.H., *et al.* (2013) Thermo-Economic Multi-Objective Optimization of Solar Dish-Stirling Engine by Implementing Evolutionary Algorithm. *Energy Conversion and Management*, **73**, 370-380.
- [2] Ponsich, A., Jaimés, A.L. and Coello, C.A.C. (2013) A Survey on Multiobjective Evolutionary Algorithms for the Solution of the Portfolio Optimization Problem and Other Finance and Economics Applications. *IEEE Transactions on Evolutionary Computation*, **17**, 321-344. <http://dx.doi.org/10.1109/TEVC.2012.2196800>
- [3] Mukhopadhyay, A., *et al.* (2014) Survey of Multi-Objective Evolutionary Algorithms for Data Mining: Part II. *IEEE Transactions on Evolutionary Computation*, **18**, 20-35.
- [4] Coello, C.A.C., Van Veldhuizen, D.A. and Lamont, G.B. (2002) Evolutionary Algorithms for Solving Multi-Objective Problems. Springer, Vol. 242. <http://dx.doi.org/10.1007/978-1-4757-5184-0>
- [5] Gong, M.-G., Jiao, L.-C., Yang, D.-D. and Ma, W.-P. (2009) Evolutionary Multi-Objective Optimization Algorithms. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.503.5879>
- [6] Larranaga, P. and Lozano, J.A. (2002) Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation. Springer Science & Business Media, Vol. 2. <http://dx.doi.org/10.1007/978-1-4615-1539-5>
- [7] Khan, N., Goldberg, D.E. and Pelikan, M. (2002) Multi-Objective Bayesian Optimization Algorithm. *Urbana*, **51**, 61801.
- [8] Pelikan, M., Sastry, K. and Goldberg, D.E. (2005) Multiobjective hBOA, Clustering, and Scalability. *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation*, ACM, New York, 663-670.

- [9] Shah, R. and Reed, P. (2011) Comparative Analysis of Multiobjective Evolutionary Algorithms for Random and Correlated Instances of Multiobjective d-Dimensional Knapsack Problems. *European Journal of Operational Research*, **211**, 466-479. <http://dx.doi.org/10.1016/j.ejor.2011.01.030>
- [10] Karshenas, H., Santana, R., Bielza, C. and Larranaga, P. (2014) Multiobjective Estimation of Distribution Algorithm Based on Joint Modeling of Objectives and Variables. *IEEE Transactions on Evolutionary Computation*, **18**, 519-542.
- [11] Zhang, Q.F., Zhou, A.M. and Jin, Y.C. (2008) RM-MEDA: A Regularity Model-Based Multiobjective Estimation of Distribution Algorithm. *IEEE Transactions on Evolutionary Computation*, **12**, 41-63.
- [12] Luis Martí, Jesús García, Antonio Berlanga, Carlos A Coello Coello, José M Molina. (2011) MB-GNG: Addressing Drawbacks in Multi-Objective Optimization Estimation of Distribution Algorithms. *Operations Research Letters*, **39**, 150-154. <http://dx.doi.org/10.1016/j.orl.2011.01.002>
- [13] Costa, M. and Minisci, E. (2003) MOPED: A Multi-Objective Parzen-Based Estimation of Distribution Algorithm for Continuous Problems. *Evolutionary Multi-Criterion Optimization*. Springer, 282-294.
- [14] Cheng, R., et al. (2015) A Multiobjective Evolutionary Algorithm Using Gaussian Process-Based Inverse Modeling. *IEEE Transactions on Evolutionary Computation*, **19**, 838-856. <http://dx.doi.org/10.1109/TEVC.2015.2395073>
- [15] Zhou, L.H., Zhou, A.M., Zhang, G.X. and Shi, C. (2011) An Estimation of Distribution Algorithm Based on Nonparametric Density Estimation. *IEEE Congress on Evolutionary Computation (CEC)*, 2011. IEEE, , 1597-1604.
- [16] Lozano, J.A. (2006) Towards a New Evolutionary Computation: Advances on Estimation of Distribution Algorithms. Springer Science & Business Media, Netherlands, Vol. 192.
- [17] Pelikan, M., Sastry, K. and Cantú-Paz, E. (2007) Scalable Optimization via Probabilistic Modeling: From Algorithms to Applications. Springer, Vol. 33.
- [18] Larranaga, P., Karshenas, H., Bielza, C. and Santana, R. (2012) A Review on Probabilistic Graphical Models in Evolutionary Computation. *Journal of Heuristics*, **18**, 795-819.
- [19] Shah, R. and Reed, P. (2011) Comparative Analysis of Multiobjective Evolutionary Algorithms for Random and Correlated Instances of Multiobjective d-Dimensional Knapsack Problems. *European Journal of Operational Research*, **211**, 466-479. <http://dx.doi.org/10.1016/j.ejor.2011.01.030>
- [20] Chen, C.-H. and Chen, Y.-P. (2007) Real-Coded ECGA for Economic Dispatch. *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, ACM, New York, 1920-1927. <http://dx.doi.org/10.1145/1276958.1277343>
- [21] Bacardit, J., Stout, M., Hirst, J.D., Sastry, K., Llorà, X. and Krasnogor, N. (2007) Automated Alphabet Reduction Method with Evolutionary Algorithms for Protein Structure Prediction. *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, ACM, New York, 346-353
- [22] Rossi, C., Abderrahim, M. and Díaz, J.C. (2008) Tracking Moving Optima using Kalman-Based Predictions. *Evolutionary Computation*, **16**, 1-30. <http://dx.doi.org/10.1162/evco.2008.16.1.1>
- [23] Larranaga, P., Etxeberria, R., Lozano, J.A. and Pena, J.M. (2000) Optimization in Continuous Domains by Learning and Simulation of Gaussian Networks. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.31.3105>
- [24] Helbig, M. and Engelbrecht, A. (2015) Benchmark Functions for cec 2015 Special Session and Competition on Dynamic Multi-objective Optimization. Tech. rep., Technical Report.
- [25] Sierra, M.R. and Coello, C.A.C. (2005) Improving PSO-Based Multi-Objective Optimization Using Crowding, Mutation and I-Dominance. In: *Evolutionary Multi-Criterion Optimization*, Springer, 505-519.
- [26] Jiang, M., Ding, Y., Goertzel, B., Huang, Z., Zhou, C. and Chao, F. (2014) Improving Machine Vision via Incorporating Expectation-Maximization into Deep Spatio-Temporal Learning. *International Joint Conference on Neural Networks (IJCNN)*, Beijing, 6-11 July 2014, 1804-1811.
- [27] Jiang, M., Huang, W., Huang, Z. and Yen, G.G. (2015) Integration of Global and Local Metrics for Domain Adaptation Learning Via Dimensionality Reduction. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=7349204&tag=1
- [28] Zhan, Z.H., Liu, X.F., Gong, Y.J., Zhang, J., Chung, H.S.H. and Li, Y. (2015) Cloud Computing Resource Scheduling and a Survey of Its Evolutionary Approaches. *ACM Computing Surveys (CSUR)*, **47**, 63. <http://dx.doi.org/10.1145/2788397>
- [29] Jiang, M., Yu, Y., Liu, X., Zhang, F. and Hong, Q. (2012) Fuzzy Neural Network Based Dynamic Path Planning. *International Conference on Machine Learning and Cybernetics (ICMLC)*, Xi'an, 15-17 July 2012, Vol. 1, 326-330.
- [30] Chao, F., Chen, F., Shen, Y., He, W., Sun, Y., Wang, Z., Jiang, M., et al. (2014) Robotic Free Writing of Chinese Characters via Human-Robot Interactions. *International Journal of Humanoid Robotics*, **11**, 1450007. <http://dx.doi.org/10.1142/S0219843614500078>
- [31] Chao, F., Lee, M.H., Jiang, M. and Zhou, C. (2014) An Infant Development-Inspired Approach to Robot Hand-Eye Coordination. *International Journal of Advanced Robotic Systems*, **11**, 1-14. <http://dx.doi.org/10.5772/57555>

- [32] Jiang, M., Zhou, C. and Chen, S. (2010) Embodied Concept Formation and Reasoning via Neural-Symbolic Integration. *Neurocomputing*, **74**, 113-120. <http://dx.doi.org/10.1016/j.neucom.2009.11.052>
- [33] Jiang, M., Yu, Y., Chao, F., Shi, M. and Zhou, C. (2013) A Connectionist Model for 2-Dimensional Modal Logic. *IEEE Symposium on Computational Intelligence for Human-Like Intelligence (CIHLI)*, Singapore, 16-19 April 2013, 54-59.
- [34] Wu, Y., Jiang, M., Huang, Z., Chao, F. and Zhou, C. (2015) An NP-Complete Fragment of Fibring Logic. *Annals of Mathematics and Artificial Intelligence*, **75**, 391-417. <http://dx.doi.org/10.1007/s10472-015-9468-4>
- [35] Cai, Z., Goertzel, B., Zhou, C., Huang, D., Ke, S., Yu, G. and Jiang, M. (2013) OpenPsi: A Novel Computational Affective Model and Its Application in Video Games. *Engineering Applications of Artificial Intelligence*, **26**, 1-12. <http://dx.doi.org/10.1016/j.engappai.2012.07.013>
- [36] Cai, Z., Goertzel, B., Zhou, C., Zhang, Y., Jiang, M. and Yu, G. (2012) Dynamics of a Computational Affective Model Inspired by Dörner's Psi Theory. *Cognitive Systems Research*, **17**, 63-80. <http://dx.doi.org/10.1016/j.cogsys.2011.11.002>